FEDERAL RESERVE BANK OF ST. LOUIS

REVIEW

SEPTEMBER/OCTOBER 2004 | VOLUME 86, NUMBER 5

Review is published six times per year by the Research Division of the Federal Reserve Bank of St. Louis and may be accessed at our web address: **research.stlouisfed.org/publications/review**. Single-copy subscriptions are also available free of charge. Send requests to: Federal Reserve Bank of St. Louis, Public Affairs Department, P.O. Box 442, St. Louis, MO 63166-0442, or call (314) 444-8808 or 8809.

ISSN 0014-9187

Contents

Volume 86, Number 5

two respects. First, it tests the EH by estimating a general vector autoregression (VAR) of the long-term and short-term rates and testing the restrictions implied by the EH on the VAR using a Lagrange multiplier test. Second, the issue of stationarity of interest rates is considered. The paper not only considers the possibility that Japanese interest rates are nonstationary, but also analyzes the implications of nonstationarity for the EH.

This paper tests for the presence of asymmetric effects of monetary policy on output. The asymmetries that the authors examine are related to the size and sign of monetary policy shocks and are based on economic theory. Using M1 as the basis for measuring monetary policy shocks, they find evidence in line with previous evidence of larger real effects resulting from positive shocks than from negative shocks—although the authors cannot reject symmetry either. However, using the federal funds rate instead, a measure that is more closely related to the actual conduct of monetary policy, they find that only small negative shocks affect real aggregate activity. The results are interpreted in terms of menu-cost models.

Free Trade: Why Are Economists and Noneconomists So Far Apart?

William Poole

Free trade—are you fer it or agin it? Why? I'm sure that this audience knows that most economists support free trade policies; however, public support for these policies can be characterized as lukewarm at best and certain groups are adamantly opposed. It is not unusual to hear the following reservations expressed about trade: "Trade harms large segments of U.S. workers." "Trade degrades the environment." "Trade exploits poor countries." We have all heard these criticisms and lots of others.

Many economists, including me, try to change public attitudes by explaining the advantages of free trade in speeches and articles intended to reach a wide range of audiences. But, let's face it: We are not very successful in changing public attitudes. Why, and how can we become more persuasive? What I will explore today is the gap that separates economists from the general public.[1]

I'll first present some evidence on the gap between economists and the general public on attitudes toward trade. I'll then outline two principles that help to understand this gap and that help to frame revealing questions when studying particular disputes. Finally, I'll offer a few suggestions on closing the gap.

Before proceeding, I want to emphasize that the views I express are mine and do not necessarily reflect official positions of the Federal Reserve System. I thank my colleagues at the Federal Reserve Bank of St. Louis for their comments; Cletus Coughlin, vice president in the Research Division, was especially helpful. However, I retain full responsibility for errors.

[1] See Coughlin (2002) for additional discussion of this gap.

THE GAP

A 1990 survey of economists employed in the United States found that more than 90 percent generally agreed with the proposition that the use of tariffs and import quotas reduced the average standard of living.[2] These results are somewhat dated; however, most observers agree that "[t]he consensus among mainstream economists on the desirability of free trade remains almost universal."[3] I don't have any data to report economists' views on particular trade disputes, but am willing to offer the following assertion: In most specific cases, disinterested economists do not defend trade restriction. By "disinterested economists" I mean economists not hired by firms engaged in the particular disputes and not employed by government agencies involved in the disputes.

If fact, I suspect that disinterested economists' attitudes about specific disputes are even more lopsided in favor of free trade than the 90 percent who generally favor free trade policies. The reason is that specific disputes almost always involve in a pretty obvious way special favors to particular industries. In contrast, economists' attitudes in general are influenced by theoretical cases in which protection may make some sense. I do not want to try to explain these theoretical cases here, but do want to note that actual trade disputes rarely fit such cases.

Let's now consider attitudes held by the general public. Public opinion polls reveal that the attitude of the general public toward free trade is not simply one of either being for free trade or for protection-

[2] See Alston, Kearl, and Vaughan (1992).

[3] See Mayda and Rodrik (2001, p. 1).

William Poole is the president of the Federal Reserve Bank of St. Louis. This article was adapted from a speech of the same title presented at the Trade, Globalization and Outsourcing Conference, Reuters America Inc., New York, New York, June 15, 2004. The author thanks colleagues at the Federal Reserve Bank of St. Louis for comments; Cletus Coughlin, vice president in the Research Division, was especially helpful. The views expressed are the author's and do not necessarily reflect official positions of the Federal Reserve System.

Federal Reserve Bank of St. Louis *Review*, September/October 2004, *86*(5), pp. 1-6.

ism.[4] Questions asking about free trade in principle reveal support for free trade, albeit not as strong as economists'. However, questions asking about free trade in practice reveal strong reservations. That is, when we get to specific trade disputes, public support for free trade tends to crumble, whereas economists rarely support trade restriction in specific disputes.

A majority of Americans do support free trade in principle. A February 2000 survey by the Pew Research Center asked the following question: "In general, do you think that free trade with other countries is good or bad for the United States?" "Good" was the response of 64 percent of the respondents, while "bad" was the response of 27 percent of the respondents. The remaining 9 percent "did not know." The general public's support for free trade is, therefore, a good bit lower than economists' support.

Much evidence exists suggesting that the general public understands the benefits from free trade in terms of increased product selection, higher quality, and lower prices. The Pew Research Center found that 81 percent of the respondents said that it was either "very good" or "somewhat good" that trade makes available different products from different parts of the world.[5]

Despite an intuitive understanding of many of the benefits of free trade, the general public has strong reservations about embracing such a policy. One set of reservations concerns distributional effects of trade. Workers are not seen as benefiting from trade. Strong evidence exists indicating a perception that the benefits of trade flow to businesses and the wealthy, rather than to workers, and to those abroad rather than to those in the United States. A poll taken by the Gallup Organization in November 1999 found that 56 percent believed that increased trade helped American companies, but that only 35 percent believed that increased trade helped American workers. In fact, 59 percent believed that trade hurts American workers.

Related to concern about adverse distributional effects of trade is the view that trade is disruptive. Regardless of whether a sufficient number of new jobs are created to compensate for the jobs lost, many Americans are reluctant to support free trade because trade causes painful adjustments for those who lose their jobs even if they find new jobs relatively quickly. The costs incurred by these workers are not necessarily offset by the creation of new and possibly better jobs.[6]

Especially noteworthy is that the sentiments of poll respondents likely reflect altruism rather than self-interest. First, only a small minority of Americans perceive the effects of trade on themselves to be negative. Second, Americans tend to view others as more vulnerable to increasing trade than themselves. Thus, it appears that the concern about the disruptive effects of job loss is for others rather than for themselves.

The concern for workers appears to go beyond U.S. borders. Based on a June 2002 survey conducted by the Chicago Council on Foreign Relations, it is clear that the majority of respondents—93 percent to be exact—think that member countries in international trade agreements should be required to maintain minimum standards for working conditions. Both moral concerns for the foreign workers and economic concerns for U.S. workers appear to affect the respondents' views.

Roughly three-quarters of the respondents to an October 1999 survey by the Program on International Policy Attitudes felt that the United States has a moral obligation to attempt to ensure that workers in foreign countries making goods for the United States do not work in harsh or unsafe conditions. Only 23 percent of the respondents felt that the United States should not judge what working conditions should be in another country. A country's national sovereignty was not viewed as a compelling reason to remain silent. Moreover, the possibility that trade expansion might improve working conditions abroad, even if not to the point of matching

4 A wealth of information on trade opinions can be found at the following website maintained by the Program on International Policy Attitudes: www.americans-world.org/digest/global_issues/intertrade/trade-general.cfm.

5 Other polls find similar results. EPIC-MRA—a polling firm conducting educational, political, industrial, and consumer market research analysis—found large majorities agreeing that trade allows American consumers to have a larger selection of goods to choose from (87 percent), improves the quality of American goods (80 percent), and allows low-income families to buy many products that they might not otherwise afford (74 percent). Polling by EPIC-MRA also found that Americans expected that they would either be paying much more (24 percent) or somewhat more (37 percent) if they were able to buy only American-made goods.

6 An October 1999 survey conducted by the Program on International Policy Attitudes asked respondents to choose between the following two statements. First: "Even if the new jobs that come from freer trade pay higher wages, overall it is not worth all the disruption of people losing their jobs." Second: "It is better to have the higher paying jobs, and the people who lost their jobs can eventually find new ones." The first statement was favored by 56 percent of the respondents, while 40 percent favored the latter statement.

conditions in the United States, was either not considered or ignored.

Additional results reveal a perception that countries that do not maintain minimum standards for working conditions have an unfair advantage that allows for the exploitation of workers and the production of goods at unduly low cost. Here there is concern about the jobs of American workers competing with cheap imports. A related aspect of this argument is that the respondents were not convinced by arguments that forcing higher standards for working conditions in foreign countries might cause elimination of jobs of extremely poor people abroad who desperately need jobs.

Strong support exists for including standards dealing with workplace health and safety, limitations on child labor, the right to strike, the right to bargain collectively, and minimum wages in trade agreements. In addition, contrary to World Trade Organization principles, Americans support unilateral decisions to bar the import of products made under substandard working conditions.

Besides the effects of increased trade on workers, many Americans are concerned that trade adversely affects the environment and that environmental standards should be incorporated into trade agreements. In a June 2002 poll by the Chicago Council on Foreign Relations, 94 percent of the respondents felt that member countries in international trade agreements should be required to maintain minimum standards for protecting the environment.[7] Support also exists for restricting the importation of goods whose production damages the environment.[8]

[7] Additional evidence supporting environmental standards in the context of trade can be found in the results of a November 2000 poll by Tarrance Group and Greenberg Quinlan Research. Respondents were asked to choose which of the following two statements were closer to their views. First: “Future trade agreements should contain safeguards that require the United States and other countries to enforce strong environmental protections, even if it limits trade.” Second: “Expanding trade is critical to the U.S. economy and trade agreements are good for our economy, even if they do not contain strong environmental protections.” The majority of respondents, 62 percent, chose the first statement as more closely reflecting their views, while only 22 percent supported not linking trade and the environment in trade agreements.

[8] An October 1999 Program on International Policy Attitudes survey asked respondents which of the following statements they agreed with the most. First: “Countries should be able to restrict the import of products if they are produced in a way that damages the environment, because protecting the environment is at least as important as trade.” Second: “If countries can put up trade barriers against a product any time they can come up with something they do not like about how it is produced, pretty soon they will be putting up barriers right and left. This will hurt the global economy and cost jobs.” Overwhelming support was found for the first statement, with 74 percent of the respondents preferring the first statement, while 22 percent supported the second statement.

On all these issues of protecting the environment, health and safety, wages and hours, working conditions, and so forth, I suspect that poll results reflect general concerns more than trade concerns per se. In the absence of a specific setting that makes the costs clear, respondents are not likely to favor accepting weaker protections for the environment, for example. Few Americans favor a world trading system in which U.S. policies on environmental and other conditions could be controlled by foreign governments through their willingness to accept goods exported by the United States. Nevertheless, these frequently expressed sentiments indicating a desire to apply U.S. standards to foreign producers do affect U.S. positions in trade disputes.

WHY THE GAP? THE SIMULTANEITY PRINCIPLE

Two principles, I believe, explain the gap between the economist’s view and the public’s view on trade. These are what I will call the “simultaneity principle” and the “political-favors principle.” I’ll discuss the first of these now and the second shortly.

The not very insightful or artful term “simultaneity principle” encompasses the economist’s case for free trade. I’m using the term because economists think about the economy through a model in which outcomes in markets are determined together as a consequence of the interactions among markets. Such interactions are represented abstractly in a mathematical model with many equations that must be solved simultaneously.

“Simultaneity principle” sounds complicated, and is meant to. I used to teach the introductory macro course to economics majors and remember well my struggle to explain the characteristics of the basic Keynesian macro model with 10 equations that had to be solved simultaneously. Teaching this material required many hours of classroom time. I could use the model to explain why, for example, an effort by households to increase their saving might have as the primary effect for a time a reduction in total employment, with the precise outcome depending on the nature of monetary policy and the degree of price flexibility. Many other exercises explain counterintuitive outcomes—counterintuitive, that is, until you have worked with the model long enough to change your intuition. It is simply a fact that the outcomes can be complicated to explain when everything in the economy depends on everything else. Indeed, in large models with scores of equations it can be difficult even for economists to identify remote and indirect effects. Elaborate simulation investigations are typically required when the models are large and complex.

The economist's case for free trade rests primarily on the fact that imposing or removing trade restrictions invariably helps some firms and people and hurts others but with a positive net benefit for the country as a whole from moving toward freer trade. As I emphasized in a speech in November 2003, a key reason why the general public is reluctant to embrace free trade is that many do not understand the benefits.[9] And the reason people do not understand the benefits is that they do not understand the interactions and connections across markets. For one example, people may see the genuine costs imposed on workers who lose their jobs to imports, but fail to see the benefits to consumers of lower-priced goods from abroad.

Economists are trained from their first course in the subject to understand the interactions across markets. The interactions are numerous and sometimes remote from the initial disturbance that sets off a chain of such interactions. It is usually possible to explain the nature of these effects to noneconomists, and formal statistical studies can often yield estimates of the magnitude of effects.

Sometimes, an interaction is pretty obvious and it may not be difficult to convey the point. For example, restricting imports of a raw material will have positive effects on domestic *producers* of the raw material, and their employees, but will hurt domestic *users* of the raw material. Indeed, by forcing up the price of the raw material, domestic producers of the finished product may find themselves at a competitive disadvantage to foreign companies with a cheaper source of the raw material. Thus, saving jobs in the industry producing the raw material comes at the cost of reduced jobs in industries using the raw material and higher costs to consumers of the finished product.

Most journalists want to smoke out all sides of a story. In the case of a story involving a trade dispute, smoking out the indirect effects is critical to explaining all sides of the story. Understanding the simultaneity principle leads immediately to questions about possible indirect and remote effects of trade restrictions. Those questions need to be addressed to economists and industry experts who can uncover the connections across markets and the indirect effects of trade restrictions.

It is important to recognize that the case for free international trade is really part of a more general case for free markets. The analysis of interregional trade within a country is in most respects exactly the same as the analysis of international trade. International trade is a separate subject within economics primarily because it deals with restrictions on trade that do not ordinarily exist between regions of a country.

Economic restrictions are of two sorts—restrictions on trade in goods and services and restrictions on movement of factors of production. In today's world, the most severe of these restrictions is on the movement of labor. Migration across national borders is controlled almost everywhere, and capital mobility is in many cases subject to some degree of restriction.

Although trade is generally free across state borders within the United States, some restrictions do exist. In making the case for free international trade, it is sometimes helpful to refer to analogies created by restrictions within the United States. One example is state professional licensing requirements that prevent doctors, lawyers, and barbers from practicing in states where they are not licensed. Another is regulation of taxis, which may prevent taxis licensed in one jurisdiction from picking up passengers at airports in other jurisdictions. This restriction creates the inefficiency of a taxicab going one way empty, even when potential passengers are waiting in a long line for a taxi. Such examples can be multiplied many times over, and are often useful in explaining the nature of inefficiencies created by trade restrictions.

One of the most difficult interactions to explain is the connection between imports and exports. Even though a country can attract capital for a time—perhaps for a period measured in decades—in the long run, imports must be paid for by exports. Most people understand this point, but not the same point put the other way—exports require imports. Restrictions on certain imports lead, quickly or eventually, either to increases in other imports or decreases in exports. This point is extremely important, for it means that "saving jobs" by restricting imports saves only jobs in the particular protected industry. Saving such jobs necessarily means losing jobs in other import-competing industries or in export industries.

Consequently, one of the points economists emphasize over and over is that saving jobs in partic-

[9] The speech was presented to the Louisville Society of Financial Analysts in Louisville, Kentucky, on November 19, 2003. It was published in the Federal Reserve Bank of St. Louis *Review*, March/April 2004, *86*(2), pp. 1-7.

ular industries does not save employment for the economy as a whole. Economists are sometimes charged with insensitivity over job losses, when in fact most of us are extremely sensitive to such losses. What good economics tells us is that saving jobs in one industry does not save jobs in the economy as a whole. We urge people to be as sensitive to the jobs indirectly lost as a consequence of trade restriction as to those lost as a consequence of changing trade patterns. Indirect job loss is part of the story of trade restriction and can be smoked out if journalists will consult knowledgeable experts.

I've already emphasized that the case for free trade is really part of the case for free, competitive markets more generally. This fact opens up another avenue for informative coverage of trade issues. Why should we be more concerned about job losses from international trade than we are about job losses from domestic competition or changing technology? Outsourcing has been an issue recently. Some firms have replaced staff handling phone inquiries with staff abroad; other firms have replaced call-center staff with automated message systems. Is it better for the caller to be able to talk with a person, who may be abroad, or to go through endless menus of the form, "press 1 if you are a retail customer, press 2 if you are a wholesale customer, press 3 if …"? When I go through these menus, I'm usually looking for "press 4 to transfer to our competitor."

WHY THE GAP? THE POLITICAL-FAVORS PRINCIPLE

Trade restriction requires legislative intervention, or regulatory intervention authorized by legislation. That means that trade restriction is inherently political. I do not mean to use "political" in a pejorative sense, for politics is an essential part of democracy and democracy is an essential part of liberty.

Legislation involving economic issues typically creates gains for some and losses for others. Every legislator is aware of this fact. Legislation is typically drawn in such a way to minimize the visibility of the losses, to avoid creating resistance to the legislation and lost votes. Legislation is often drawn to increase the visibility of the gains to those who benefit, to attract votes. However, sometimes legislation hides the benefits, to reduce the possibility that publicity will lead to opposition. Those who benefit, of course, may be well aware of the benefit. It is perfectly natural that legislators should write legislation this way. You and I would do the same thing if we were legislators.

Because they understand the importance of the political-favors principle, journalists know immediately what sorts of questions to ask. When evaluating a particular trade restriction, who gains and who loses? What is the net for the economy as a whole of the gains and losses?

When I read a story that reports only the benefits of trade restriction, I know the story is incomplete. I also know that losers from restriction often do not realize they've been hurt. I'm reminded of the story some years ago of a bank employee who found a way to skim fractional interest payments into his own account. The depositor didn't realize that his account had been rounded off to $308.27 whereas his account really had $308.274. The extra 4 tenths of a cent, if left in the account, would have earned interest and have led to a larger account balance in the future. The accountant who skimmed a few tenths of a cent from thousands of accounts put a lot into his own account, until he got caught. Many trade restrictions work this way—they cost consumers just a little, but add up to a lot for the protected industry.

Perhaps there is no reason to feel much outrage about such trade restriction, but in most cases legislators would not be able to impose a small sales tax on the good and funnel the revenues to the favored industry. The stratagem works when it is hidden. Telling the full story of any particular trade restriction may require adding up lots of pennies extracted from those who do not realize they are paying.

CLOSING THE GAP

Once the reasons for the gap between economists and noneconomists are understood, approaches to closing the gap become clear. I've already emphasized the important role of journalists. In this area, as with all other public policy areas in a democracy, a free and enterprising press is essential to effective government in the interest of the nation at large.

As a former university professor, it is natural for me to believe that formal education plays an important role. Nevertheless, every educator is aware of the short half-life of much of the material taught. Students' knowledge usually peaks at exam time, and then starts to decay. What I hope my students retained is some very basic principles, such as the gains from voluntary exchange, and respect for economics as a discipline. Years after formal study, people need to be reminded of the analysis and how it applies to real-life policy issues. Educators can play a continuing role, by writing and speaking for noneconomist audiences.

But I began this speech by expressing disappointment over the effectiveness of economists' speeches, and that is why I'm emphasizing the importance of the role of the press today. I've suggested that every story on trade issues, to be complete, must explore who gains, who loses, and the net of gains and losses for the nation as a whole.

Whenever faced with a policy choice that creates winners and losers, we face the difficult problem of somehow weighting one person's benefit against another's loss. The issue appears constantly, and we take two general approaches. The first is that the government does not take property without compensation. The second is that the government stands aside from the competitive market system and lets the chips fall as they may.

Government provides compensation when it takes land for a highway. It is important to note, however, that the compensation is an estimate of fair market value. We understand the loss to a family when government takes land that has been in the family for generations, but we do not try to compensate for the sentimental value of the land. It is simply not possible to maintain a vigorous, growing economy while giving great weight and actual compensation for loss of sentimental value.

Government provides generalized compensation, or adjustment assistance, through unemployment insurance. The United States does not have a general program to compensate owners of capital. Unemployment assistance is relatively limited, as it must be to retain incentives to return to work. Existing legislation also provides some extra benefits for adjustment to losses arising from international trade. My view of this legislation is that in the abstract there is no particular reason to provide more assistance for job loss due to international trade than for any other reason, but as a practical matter such assistance is warranted if it helps to gain acceptance for trade liberalization. We should recognize that many of the arguments for maintaining certain industries in the United States are essentially sentimental, the case being essentially the same as that for avoiding taking land that has been the family farm for generations.

We live in a society that on the whole accepts an economic system that lets the chips fall where they may. Some decry the nature of this system, but its general support rests on the progress and the higher standard of living it affords. We should not underestimate the individual protections built into this system. Our sophisticated market system includes insurance markets that permit individuals and firms to protect themselves against many forms of risk. More importantly, the vitality of our markets creates opportunities for new firms and new employment to absorb those displaced by changing competitive conditions. Our dynamic economic system, and not restrictive trade legislation, provides the best protection for our citizens.

THE BOTTOM LINE

We all know that a vigorous and just democracy depends on a free and enterprising press. I urge you to keep my two principles—the simultaneity principle and the political-favors principle—in mind when reporting on trade issues. The first requires that you identify the complicated and indirect effects of trade restrictions, and the second requires that you understand the winners and losers from restrictions. I believe that the general voting public will be more likely to favor free trade policies if it understands the issues at a deeper level.

So remember: Every trade story requires at least three sections. One reports who gains, one reports who loses, and one reports the net of the gains and losses for the country as a whole. There is an enormous opportunity here: Sound and impartial reporting case by case by case will do more, I believe, to promote free trade policies than all the economists' speeches extolling the benefits of trade laid end to end.

REFERENCES

Alston, Richard M.; Kearl, J.R. and Vaughan, Michael B. "Is There a Consensus among Economists in the 1990's?" *American Economic Review*, May 1992, *82*(2), pp. 203-29.

Coughlin, Cletus C. "The Controversy Over Free Trade: The Gap Between Economists and the General Public." Federal Reserve Bank of St. Louis *Review*, January/February 2002, *84*(1), pp. 1-22.

Mayda, Anna Maria and Rodrik, Dani. "Why Are Some People (and Countries) More Protectionist than Others?" Working Paper No. 8461, National Bureau of Economic Research, September 2001.

Poole, William. "A Perspective on U.S. International Trade." Federal Reserve Bank of St. Louis *Review*, March/April 2004, *86*(2), pp. 1-7.

Trends in the Efficiency of Federal Reserve Check Processing Operations

David C. Wheelock and Paul W. Wilson

In 2000, U.S. consumers, businesses, and government entities made some 71.5 billion non-cash payments, with a value of approximately $46.6 trillion. Paper checks accounted for about 60 percent of the total number of non-cash retail payments, with credit and debit cards and automatic clearinghouse (ACH) payments making up the remainder. Approximately 29 percent of checks were deposited or cashed at the same depository institution on which they were drawn, so-called "on us" checks. Of the remaining 71 percent, a high percentage were processed by the Federal Reserve System, moving physically through one or more Fed check processing facilities, which are located in Federal Reserve Banks, Branches, and dedicated processing offices.[1] In 2002, the Federal Reserve processed some 16.6 billion checks, generating $759 million in revenue and $744.3 million in expenses for the System (Board of Governors of the Federal Reserve System, 2002, pp. 128, 139).

Both the number of checks written and the number processed by the Federal Reserve have been declining since the late 1990s, as payments are increasingly made electronically. Survey data show a decline in the number of checks paid from 49.5 billion in 1995 to 42.5 billion in 2000, the last year for which data are available, while the number of electronic payments increased by 14.2 billion items. Between 1995 and 2000, the share of total non-cash payments made by checks declined from 77 percent to 60 percent (Gerdes and Walton, 2002). The number of checks cleared by the Federal Reserve has also been declining. The Fed's processing volume peaked in 1992 at over 19 billion checks. New settlement rules adopted in 1994 caused the number of checks processed by the Fed to fall sharply, to a low of 15.5 billion in 1995.[2] Although the Fed's processing volume rose annually between 1995 and 1999, in 1999 the Fed handled 17.1 billion checks and volume has since been declining steadily.

The Monetary Control Act of 1980 requires the Federal Reserve to charge fees for providing payments services, including check processing, to cover (i) the Fed's expenses and (ii) imputed taxes and profits that would be earned by private firms providing similar services. Because of declining volume, the Federal Reserve recently determined that it must reduce the number of its offices that process checks to remain in compliance with the cost-recovery requirements of the Monetary Control Act.

Legislation taking effect in October 2004 has the potential to reduce further the volume of paper checks processed by the Federal Reserve. The Check Clearing for the 21st Century Act, or Check 21, creates a new negotiable instrument, called a substitute check. Substitute checks are printed reproductions made from digital images of the original paper checks and are the legal equivalent of the original checks. The legal status of substitute checks will facilitate the truncation of checks and greater use of electronic check processing. It will thus replace some of the physical movement of paper checks from the banks where checks are cashed or deposited to the banks on which they are drawn or returned.

[1] These data are based on comprehensive surveys of payments activity sponsored by the Federal Reserve. See Gerdes and Walton (2002).

[2] Rules were implemented through the Fed's Regulation CC on January 3, 1994, that enabled collecting banks to receive same-day settlement by presenting checks directly to paying banks by 8:00 a.m. Checks cleared by the Fed declined 13.3 percent in 1994, as banks found it advantageous under the new rules to switch some clearing of checks to private clearinghouses.

David C. Wheelock is an assistant vice president and economist at the Federal Reserve Bank of St. Louis. Paul W. Wilson is a professor of economics at the University of Texas at Austin; this research was carried out while Wilson was a visiting scholar at the Federal Reserve Bank of St. Louis. The authors thank Tom Garrett and Alton Gilbert for comments on a previous version of this article. Heidi L. Beyer, Joshua Ulrich, and Neil Wiggins provided research assistance.

Federal Reserve Bank of St. Louis *Review*, September/October 2004, *86*(5), pp. 7-19.

The Monetary Control Act sought to use market discipline to improve the efficiency with which the Fed provides payments services. If the demand for its paper check processing continues to fall, the Fed will face intensified pressures to reduce its costs if it is to remain in compliance with the Act. This article investigates patterns in the efficiency of Fed check processing operations over time, and it introduces a new technique for estimating efficiency that overcomes estimation problems inherent in methods used previously to examine the efficiency of Fed payments services.

The next section describes the efficiency concept used in this article. We then describe the processing of checks by Federal Reserve offices and discuss our empirical model and data. Subsequent sections discuss alternative ways of estimating efficiency, focusing on the nonparametric data envelopment analysis (DEA) and order-*m* estimators, and present estimation results and conclusions.

INPUT TECHNICAL EFFICIENCY

This article examines changes in the technical efficiency of Federal Reserve check offices over time. A decisionmaking unit (a Fed check office in our context) is input-inefficient if it uses more input to produce a given quantity of output than the existing technology requires. For production processes involving a single input, input-technical inefficiency is measured simply as the ratio of input actually used to the minimum feasible input amount required to produce a given amount of output. For processes involving multiple inputs, inefficiency is measured by the proportionate overuse of all inputs.

Inefficiency can also be measured by the extent to which an office produces less than the technically feasible amount of output from given amounts of input. The output of a Fed check office (e.g., the number of checks processed) is largely outside the office's control, however, at least in the short run. Hence, we focus on *input* technical efficiency, rather than *output* technical efficiency. Although technical efficiency is among the most frequently investigated types of efficiency, a comprehensive analysis of Fed check offices, or any other type of organization, would require examination of a variety of performance measures.[3]

Thus far, evidence on the impact of the Monetary Control Act on the Fed's efficiency in providing payments services has been mixed. That evidence, which is discussed more fully in Gilbert, Wheelock, and Wilson (2004), suggests that the Monetary Control Act did not result in an immediate improvement in the efficiency with which the Fed delivers services. Gilbert, Wheelock, and Wilson (2004) find, however, that Fed offices generally became more *productive* during the 1990s, with considerable convergence across offices. Nevertheless, that study also finds that as of 1999, the median Fed office could have feasibly reduced its input usage by about 30 percent without reducing output.[4]

Although Gilbert, Wheelock, and Wilson (2004) find that, over time, Fed offices increased output (e.g., checks processed) for given amounts of input (e.g., labor and processing equipment), this increase in productivity does not imply that Fed offices necessarily became more efficient; that is, they did not necessarily move closer to the production frontier. The extent to which an office is (in)efficient is reflected in the difference between the amount of output the office actually produces and the amount it could feasibly produce using available technology. Technological improvement implies an increase in the amount of output that can be produced for given amounts of input, i.e., a shift in the production frontier. Therefore, because of technological improvement, an office could become more productive, i.e., produce more output using given amounts of input, without becoming more efficient, i.e., moving closer to the production frontier. In this article, we control for changes in (estimated) technology to determine whether Fed offices generally became more efficient in terms of feasible production over time. At the same time, we use newly developed techniques for estimating efficiency and a longer sample period than Gilbert, Wheelock, and Wilson (2004).

---

[3] Other measures of efficiency include allocative efficiency, which takes account of input prices to determine the efficient mix of inputs, and scale efficiency, which refers to operation at efficient scale. Other common efficiency measures include cost efficiency, which examines the extent to which a firm minimizes total cost, given input prices and output quantities, and profit efficiency, which examines the extent to which a firm maximizes profit.

[4] Estimates of average inefficiency obtained by this and other studies of Fed check production range from less than 5 percent to 35 percent, depending on the type of efficiency studied, estimator used, and sample period. Such estimates are in line with estimates of inefficiency for private sector firms, including commercial banks. See Berger and Humphrey (1997) for a survey of research examining inefficiency among private sector firms.

Table 1

Definitions and Measurement of Inputs

1. **Personnel**: number of employee hours
2. **Materials**, **Software**, **Equipment**, and **Support**

   Expenditures are deflated by the following price measures to obtain physical units, which are then combined using a Tornquist index:

   Materials: GDP implicit price deflator (seasonally adjusted, 1996 = 100)

   Software: Private nonresidential fixed investment deflator for software (seasonally adjusted, 1996 = 100)

   Equipment:

   For 1979-89: PPI for check-handling machines (June 1985 = 100)

   For 1990-2003: PPI for the net output of office machinery manufacturing (not seasonally adjusted, June 1985 = 100)

   Support: GDP implicit price deflator (seasonally adjusted, 1996 = 100)
3. **Transit** (Shipping, Travel, Communications, and Data Communications Support)

   Expenditures are deflated by the following price measures to obtain physical units, which are then combined using a Tornquist index:

   Shipping and *Travel:* Private nonresidential fixed investment deflator for aircraft (seasonally adjusted, 1996 = 100)

   Communications and *Data Communications Support:* Private nonresidential fixed investment deflator for communications equipment (seasonally adjusted, 1996 = 100)
4. **Facilities***:* Expenditures are deflated by the following price index: “Historical Cost Index” from *Means Square Foot Costs Data 2000* (R.S. Means Company: Kingston, MA, pp. 436-42). Data are January values.

SOURCE: Federal Reserve Planning and Control System documents unless otherwise noted.

Federal Reserve Check Processing

The clearing of checks involves receiving checks from depositing banks (defined broadly to include all depository institutions), sorting them, crediting the accounts of depositing banks, and delivering the checks to the banks upon which they are drawn. Such “forward items” processing is the main source of revenue and total cost for Fed check operations. Some Fed offices process federal government checks and postal money orders, as well as commercial checks. Fed offices also process “return items” (which include checks returned because of insufficient funds) and “adjustment items” (which arise because of processing or other errors) and provide various electronic check services, such as imaging and truncation. Following other studies, we focus on the forward processing of commercial and federal government checks.

The methods we use permit estimation of the efficiency of check offices with multiple outputs. In addition to the number of forward items processed, we also treat the number of endpoints served by a Fed check office as an output. An endpoint is an office of a depository institution to which the Fed delivers checks; hence, the number of endpoints is a measure of the level of service provided by a Fed office. Presumably, an office serving many endpoints provides a higher level of service than an office serving few endpoints. In this sense, check processing is analogous to the delivery of mail by a post office. The output of a post office is not simply the number of items it delivers, but also the number of addresses to which it delivers mail. A post office that delivers mail to a single address provides a lower level of service than an office that delivers the same quantity of mail to several addresses.[5]

Federal Reserve offices incur a variety of costs associated with the processing of checks. Estimation of efficiency using statistical methods requires the specification of a model of the production process with a limited number of inputs. We follow Gilbert,

[5] Gilbert, Wheelock, and Wilson (2004) provide statistical evidence that the number of endpoints should be treated as a distinct, second output of check processing.

Table 2

Summary Statistics for Inputs and Outputs

	Mean	Median	Variance	Minimum	Maximum
1980:Q1–1994:Q4					
Checks (1000s)	80,734	69,985	2.25e9	10,413	265,631
Endpoints	430	366	9.33e4	32	1,686
Personnel (hours)	35,686	28,741	6.63e8	4,905	201,529
Material, etc.	1,899	1,567	1.30e6	155	7,403
Transit	1,509	1,328	8.14e5	148	6,678
Facilities	716	485	3.62e5	63	4,438
1995:Q1–2000:Q4					
Checks (1000s)	93,267	85,324	2.19e9	17,205	280,006
Endpoints	351	339	4.21e4	32	1,262
Personnel (hours)	30,827	26,825	3.41e8	5,478	111,204
Material, etc.	2,008	1,658	1.82e6	373	10,630
Transit	484	372	1.26e5	20	2,124
Facilities	741	561	2.94e5	129	3,991
2001:Q1–2003:Q3					
Checks (1000s)	104,579	96,559	2.70e9	18,253	292,891
Endpoints	314	290	2.85e4	92	955
Personnel (hours)	27,931	22,930	2.57e8	4,357	111,497
Material, etc.	2,597	2,303	2.11e6	424	9,853
Transit	516	403	1.14e5	111	1,805
Facilities	935	780	3.98e5	160	3,870

Wheelock, and Wilson (2004), which in turn follows Bauer and Hancock (1993) and Bauer and Ferrier (1996), in defining four distinct categories of inputs used in the processing of forward items and serving endpoints: (i) personnel; (ii) materials, software, equipment, and support; (iii) transit; and (iv) facilities. Our model requires estimates of the physical quantities used of each input rather than total expenditures. Table 1 describes how we construct each of the four inputs using expense data for each Fed check facility; Table 2 provides summary statistics for the four inputs and two outputs. Our data include quarterly observations spanning 1980:Q1–2003:Q3 for each Fed office that operated at any time during this period. Because of changes in the Fed's cost accounting system, there are discontinuities in the data at the ends of 1994 and 2000. Hence, we report summary statistics for 1980:Q1–1994:Q4, 1995:Q1–2000:Q4, and 2001:Q1–2003:Q3 separately.

ESTIMATING EFFICIENCY

Statistical estimation of efficiency requires a model relating production inputs and outputs. Many studies estimate translog cost or profit functions that include a two-sided random noise term and a one-sided random inefficiency term. The translog functional form has been shown to misspecify cost relationships for several types of firms, however, including commercial banks (see, e.g., McAllister and McManus, 1993; Wheelock and Wilson, 2001). Instead of estimating translog or other parametric functions, some studies use nonparametric methods to estimate efficiency. An estimate of inefficiency for an individual office consists of a measure comparing that office with an estimate of best practice. Common nonparametric estimators of the production frontier are the data envelopment analysis (DEA) and the free disposal hull (FDH) estimators. Nonparametric estimators do not require specification

Figure 1

True and Estimated Production Frontier

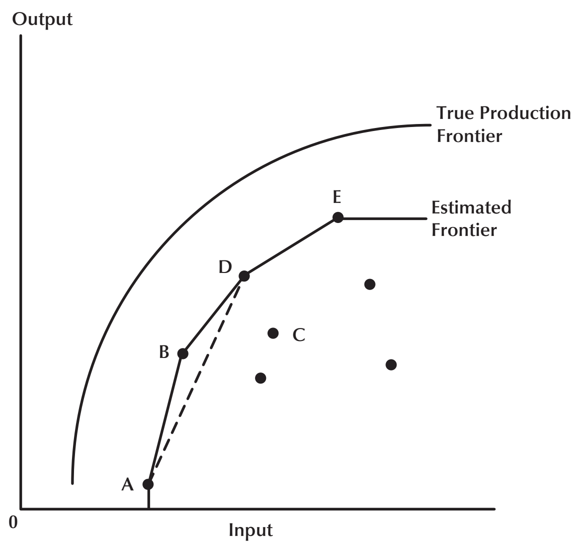
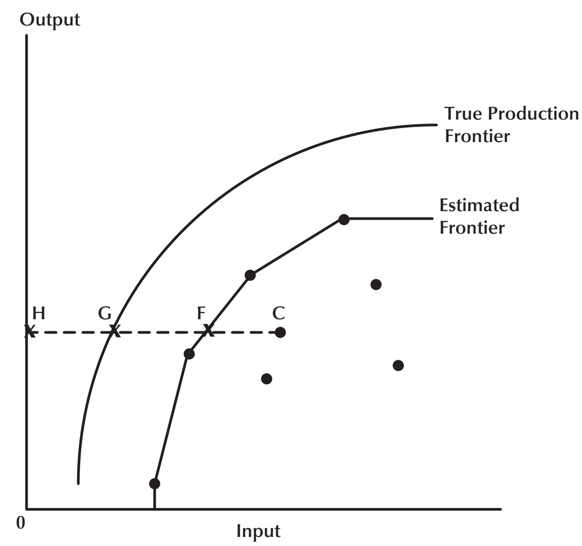


Figure 2

Estimating Efficiency



of a particular functional relationship between production inputs and outputs. DEA and FDH do impose certain assumptions about the shape of the production surface, or technology, however, and applications often require very large data sets to obtain meaningful efficiency estimates. DEA and FDH are also highly sensitive to extreme observations and noise in the data.[6]

The order-*m* estimator proposed by Cazals, Florens, and Simar (2002), by contrast, is robust to extreme values and noise. Further, the order-*m* estimator imposes fewer assumptions on the shape of the production surface than DEA. Perhaps most importantly, for large numbers of inputs and outputs, the order-*m* estimator requires far less data to obtain meaningful estimates of inefficiency than do frontier estimators, such as DEA or FDH. This section presents non-technical descriptions of DEA and order-*m* estimation. Readers interested in detailed treatments are referred to Cazals, Florens, and Simar (2002) and Wheelock and Wilson (2003).

Data Envelopment Analysis

DEA uses observations on outputs and inputs of decisionmaking units (e.g., firms or, in our application, Federal Reserve check offices) to estimate the most productive combinations of outputs and inputs that are technically feasible, i.e., "efficient" combinations. The inefficiency of specific firms can be estimated by comparing their actual input/output combinations with efficient combinations. The more input a firm uses to produce given quantities of output (or, equivalently, the less output a firm produces from given quantities of inputs) relative to an efficient combination, the less efficient is the firm.

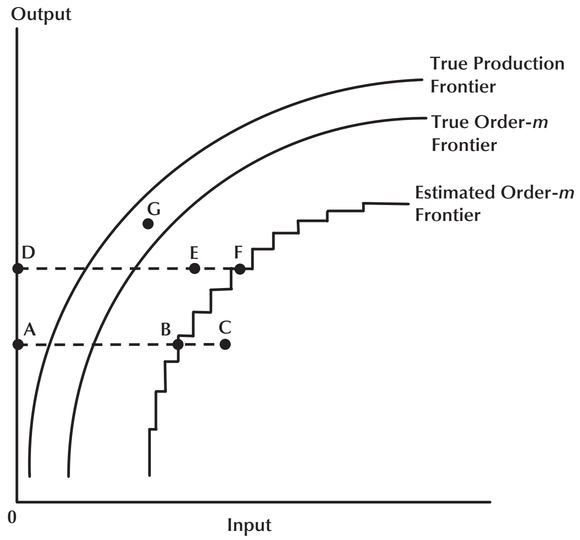
Figure 1 illustrates the technique for estimating the relationship between input and output for a production process that has one input and one output. The dots represent observed levels of input and output for eight firms. If we had complete information about the minimum level of input that firms require to produce a given level of output, we could trace out the *true* production frontier. We could then assess the degree to which a particular firm is inefficient by measuring its distance from the frontier. Given the amount of output produced, the inefficiency of a firm reflects the reduction in the input the firm would use if it were on the frontier. Alternatively, we could measure inefficiency in terms of the increase in output a firm could produce, for a given level of input, if the firm operated on the frontier.

Lacking complete information, we must assess the efficiency of firms from observations of the inputs and outputs of actual firms. We trace out the *estimated* frontier in Figure 1 by connecting the dots for firms A, B, D, and E. The only constraint we

6 See Simar and Wilson (2000) and Kneip, Simar, and Wilson (2003) for discussion of the statistical properties of FDH and DEA estimators.

Figure 3

Order-*m* Efficiency



impose on the shape of the estimated frontier is that it does not include firms like C. More precisely, we assume the frontier is convex, i.e., that any two points on or under the frontier can be connected by a line segment that never passes above the frontier. Firm C is inefficient relative to the observed frontier in the sense that it could feasibly produce the same level of output using less input (or, equivalently, produce more output from the same amount of input).

We illustrate the DEA efficiency measure in Figure 2. Consider the firm with the amounts of input and output labeled C. That firm produces an amount of output equal to the vertical distance 0H using an amount of input equal to the horizontal distance HC. If firm C were more efficient, it could use less input to produce 0H output (or, equivalently, use HC input to produce more output than 0H). The extent to which firm C could improve is reflected in the difference in the amount of input used by firm C and the horizontal distance GC. Because we do not know the location of the true frontier, however, we estimate the potential improvement in efficiency of firm C as the difference between HC and the amount of input that a firm (perhaps hypothetical) located on the estimated frontier and producing the same amount of output as firm C would use (distance FC). To measure inefficiency, we divide the distance HC by the distance HF. This ratio is greater than 1.0 for all firms that lie off the estimated frontier; and the larger this ratio, the *less* efficient the firm.[7]

Estimation of the efficient frontier requires certain assumptions. Standard assumptions, which we impose here, are that (i) the production set is convex and closed; (ii) all production requires the use of some inputs, and both inputs and outputs are strongly disposable; (iii) the observed set of inputs and outputs for check processing offices results from independent draws from a probability density function with bounded support over the production set; (iv) this density is strictly positive for all points along the frontier; (v) starting from any point along the frontier, the density is continuous in any direction toward the interior of the production set; and (vi) the true frontier is smooth. Together, these assumptions define the data-generating process that produces sample observations, and permit statistical estimation and inference about the unobserved technology as well as the unobserved input distance function.[8]

*Order-*m *Estimators*

As an alternative to nonparametric estimators of the production frontier, such as DEA, Cazals, Florens, and Simar (2002) propose estimators based on the *expected minimum input frontier* of order *m*.[9] Order-*m* estimators do not impose the assumption that the production set is convex, and in addition they permit noise (with zero expected value) in input measures. Note that DEA estimates of the production frontier can be severely distorted by extreme values. Individual observations have much less influence on order-*m* estimation. Further, for given numbers of inputs and outputs, the order-*m* estimator requires far less data to produce meaningful efficiency estimates than DEA or FDH estimators.[10]

Consider again a production process involving firms that use a single input to produce a single output. In this case, the expected minimum input frontier of order *m* is simple to estimate: (i) For each

[7] This measure of productivity is the Shephard input distance function. See Shephard (1970).

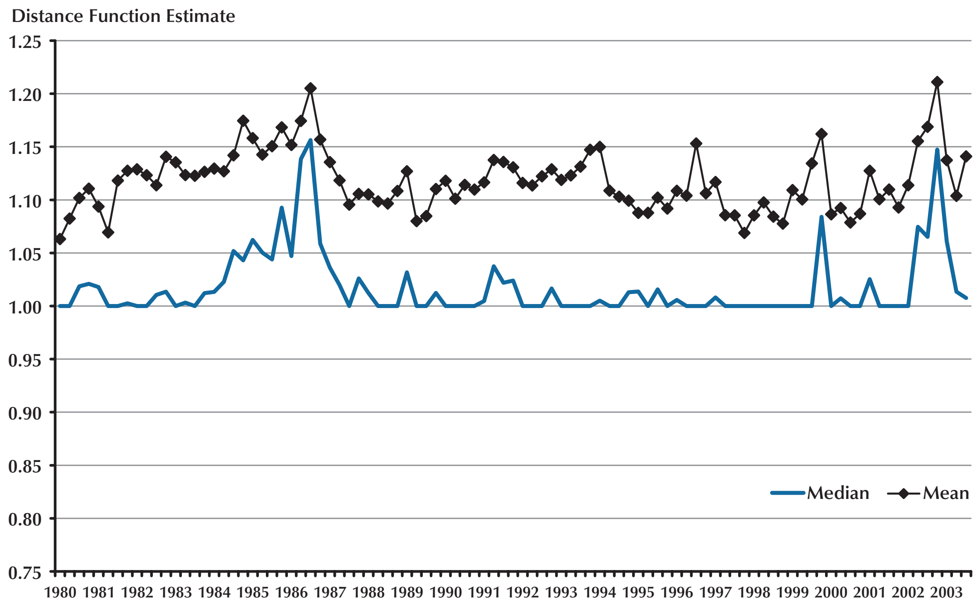
[8] See Simar and Wilson (2000) for more detailed discussion of these assumptions and the DEA estimator.

[9] Order-*m* estimators can also be applied in the output direction, i.e., to estimate the expected maximum output frontier of order *m*.

[10] Specifically, order-*m* estimators converge to their population counterparts at a rate equal to the square root of the sample size, which is far faster than the convergence rates of DEA and FDH estimators in our application. See Wheelock and Wilson (2003).

Figure 4

DEA Efficiency



firm, identify all other firms that produce at least as much output as that firm. (ii) Draw *m* times, independently, with replacement, from this set of firms, identifying the firm among the *m* draws that uses the *minimum* amount of input. (iii) Repeat step (ii) *k* times. (iv) Compute the mean input usage of the *k* firms identified in the sampling of steps (ii) and (iii). The estimated expected minimum input frontier constitutes the means computed for each sample observation.[11] The estimated frontier will lie below and to the right of the true order-*m* frontier, which itself will lie below and to the right of the true production frontier, as illustrated in Figure 3.[12]

For any given firm, an estimate of order-*m* inefficiency is obtained by computing the distance from the firm to the estimated order-*m* frontier, as illustrated in Figure 3. A hypothetical expected minimum input frontier of order *m* is traced out by the various line segments as shown. For firm C, inefficiency is estimated by the ratio AC/AB.[13] In other words, based on computation of the order-*m* frontier, firm C uses 100[(AC/AB) – 1] percent more input than the minimum amount the firm could be expected to use. Note that some firms might use less input than the expected minimum amount, as illustrated by firm E in Figure 3. For firm E, "inefficiency" is estimated by the ratio DE/DF, which is less than 1. This firm uses 100(DE/DF) percent of the expected minimum amount of input, based on order-*m* sampling of all firms that produce as much output as firm E. Finally, note that some firms, e.g., firm G, might lie above the true order-*m* frontier.[14]

ESTIMATION RESULTS

We present efficiency estimates based on both the DEA and order-*m* estimators. For each quarter, we use DEA to estimate the production frontier from observations on all offices producing positive amounts of both outputs (forward check items and

[11] The choice of *m* and *k* will depend on the particular application at hand, as discussed later.

[12] The order-*m* frontier need not be convex, as illustrated by the "stair-step" shape of the estimated frontier in Figure 3.

[13] As with DEA, this distance corresponds to the Shephard (1970) input distance function.

[14] In the case of multiple outputs, the *m* firms are drawn from the set of firms that produce at least as much of *all* outputs as the firm of interest. When there are multiple inputs, the minimum input usage among *m* firms is determined by a minimax algorithm. This minimum converges to the FDH estimate of the production frontier as *m* approaches infinity. FDH differs from DEA only in that it does not assume the production frontier to be convex. For a mathematically precise description of the order-*m* frontier in the case of multiple inputs and/or outputs, see Cazals, Florens, and Simar (2002) or Wheelock and Wilson (2003).

endpoints).[15] We then estimate the inefficiency of each office using the DEA distance function estimator described here previously.

Our approach differs from that of Gilbert, Wheelock, and Wilson (2004). Because we estimate a production frontier for every period, we obtain efficiency estimates for every office in every quarter of our sample. By contrast, Gilbert, Wheelock, and Wilson (2004) pool observations on Fed offices over time and estimate a single frontier. Consequently, they obtain efficiency estimates only for the final period of their sample (1999:Q4). Gilbert, Wheelock, and Wilson (2004) do produce distance function estimates for every office in every period relative to the single frontier, however, and changes in those estimates over time reflect changes in the productivity of individual offices. Moreover, because pooling effectively increases their sample size, the estimate of the production frontier for the final period obtained by Gilbert, Wheelock, and Wilson (2004) is probably more reliable than our quarterly estimates.

Figure 4 plots the mean and median efficiency (distance function) estimates across all offices in each quarter. DEA efficiency estimates are greater than or equal to 1, with an estimate of 1 implying that an office is fully efficient. Larger estimates imply lower efficiency.

On average, Fed offices appear to have become less efficient between 1980 and 1982, when the pricing requirements imposed by the Monetary Control Act were fully implemented. Efficiency continued to worsen to a peak in the third quarter of 1986. Mean and median efficiency then improved to approximately their pre-1982 levels. From 1987 to 2003, mean inefficiency ranged between 5 and 20 percent in all but one quarter. That is, on average, Fed offices used 5 to 20 percent more input to produce given amounts of output than DEA estimation indicates was technically feasible. Thus, our estimates suggest that Fed offices became less efficient when the Monetary Control Act was first implemented, but by the late 1980s inefficiency had stabilized. We find little evidence to suggest that the Monetary Control Act improved average efficiency in the long run, though median inefficiency was low throughout the late 1980s and 1990s before spiking in 1999 and again in 2002-03. Indeed, our estimates indicate that in several quarters more than half of Fed offices were fully efficient.[16]

As noted previously, there are reasons to be suspicious of DEA efficiency estimates, particularly in cases where the number of observations used to estimate the efficient frontier is small. The order-m class of estimators provides another way of looking at efficiency that is less sensitive to extreme-value observations and small samples.[17]

Figures 5 and 6 plot quarterly mean and median values of order-m distance estimates across all check offices for four values of m. As with our DEA estimation, for each value of m, we estimate a distinct order-m frontier for each quarter using observations on check offices for only that quarter. Recall that the order-m distance estimate for any given office in any period is obtained by comparing the actual input usage of that office with the expected minimum input usage of the office, where the expected minimum is obtained by drawing k samples of m offices that produce at least as much output as the office of interest. The minimum input usage among the m offices in each of the k samplings is recorded, with the expected minimum input usage for the office of interest calculated as the mean of the minimums observed in each of the k samplings. We set k equal to 200, which, by the law of large numbers, should be sufficient to obtain accurate estimates of the true mean without being computationally prohibitive.

The order-m distance estimate for each check office reflects the extent to which that office uses a different amount of input to produce its output level from the expected minimum input amount, as determined by the sampling algorithm described previ-

[15] Between 43 and 48 offices are used to estimate the frontier in each quarter. A few observations were dropped because of missing data.

[16] Gilbert, Wheelock, and Wilson (2004) find general improvement in the productivity of Fed check offices from the late 1980s through the 1990s. Our results, in combination with those reported by Gilbert, Wheelock, and Wilson (2004) suggest that the technology of Fed check processing was improving.

[17] Another reason to be suspicious of our DEA efficiency estimates is that they appear to be heavily dependent on the assumption that the production frontier is convex. We also used the FDH estimator, which relaxes the assumption of convexity but is otherwise identical to DEA, to estimate efficiency. Using FDH, we find that average inefficiency of Fed offices ranged from 0 to 4 percent throughout 1980-2003 and that the median inefficiency across Fed offices was always zero, i.e., that at least half of Fed offices were fully efficient in every quarter (typically, some 90 percent of offices were found to be fully efficient). Hence, our DEA estimates are strongly affected by the convexity assumption. This may not be a bad assumption, but one might not want to draw strong conclusions based on results that are so heavily influenced by a single assumption.

Figure 5

Mean Order-*m* Efficiency

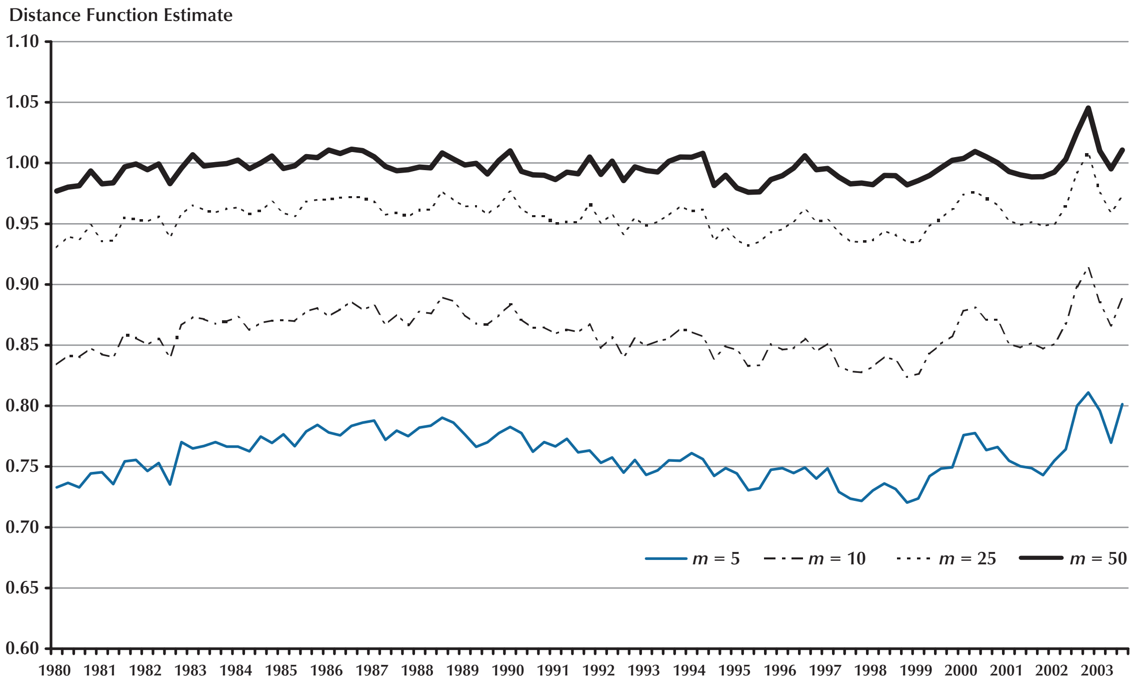


Figure 6

Median Order-*m* Efficiency

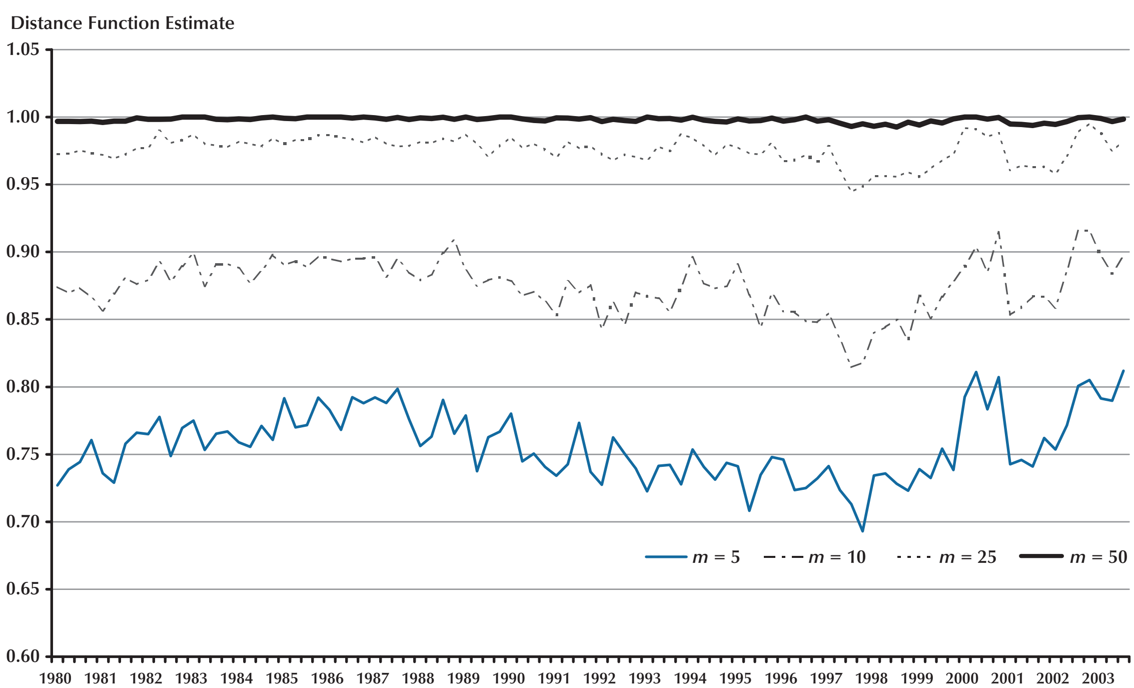


Figure 7

Order-*m* Efficiency Estimates, *m* = 5 vs. *m* = 10

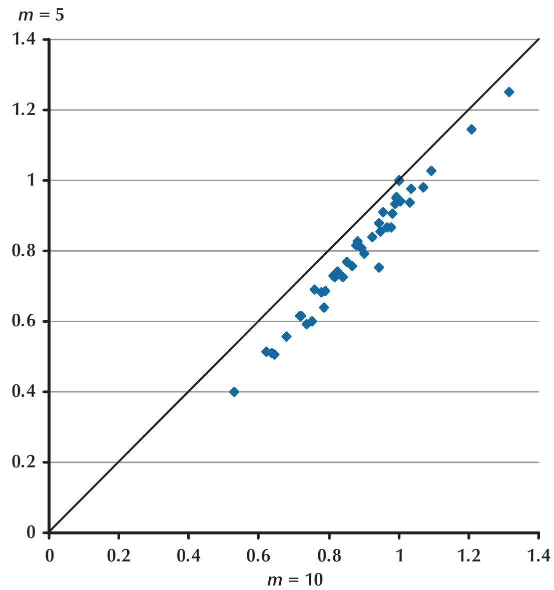


Figure 9

Order-*m* Efficiency Estimates, *m* = 25 vs. *m* = 50

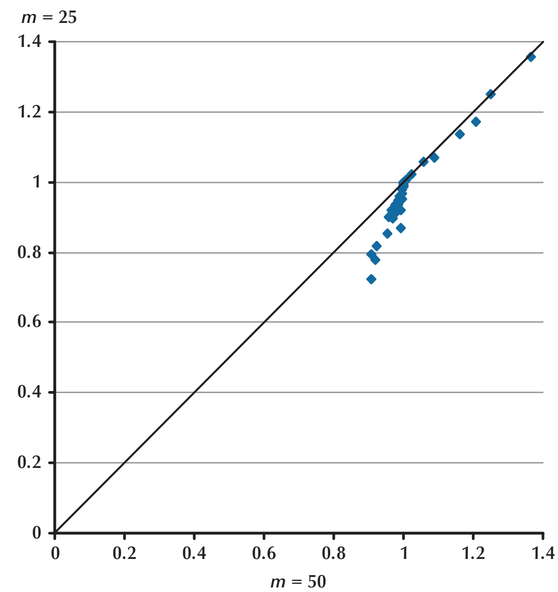
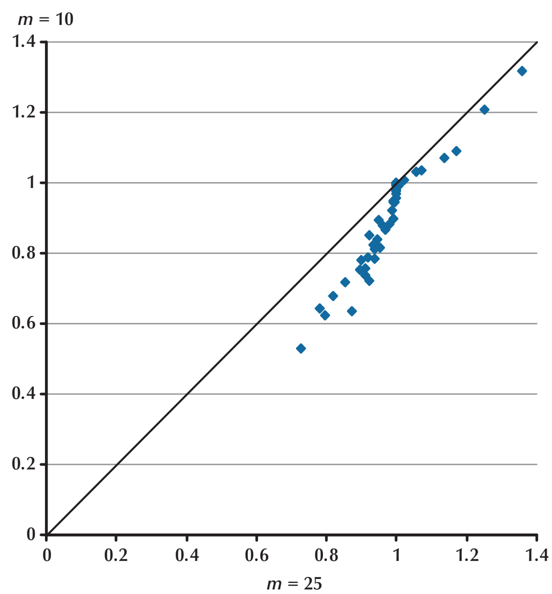


Figure 8

Order-*m* Efficiency Estimates, *m* = 10 vs. *m* = 25

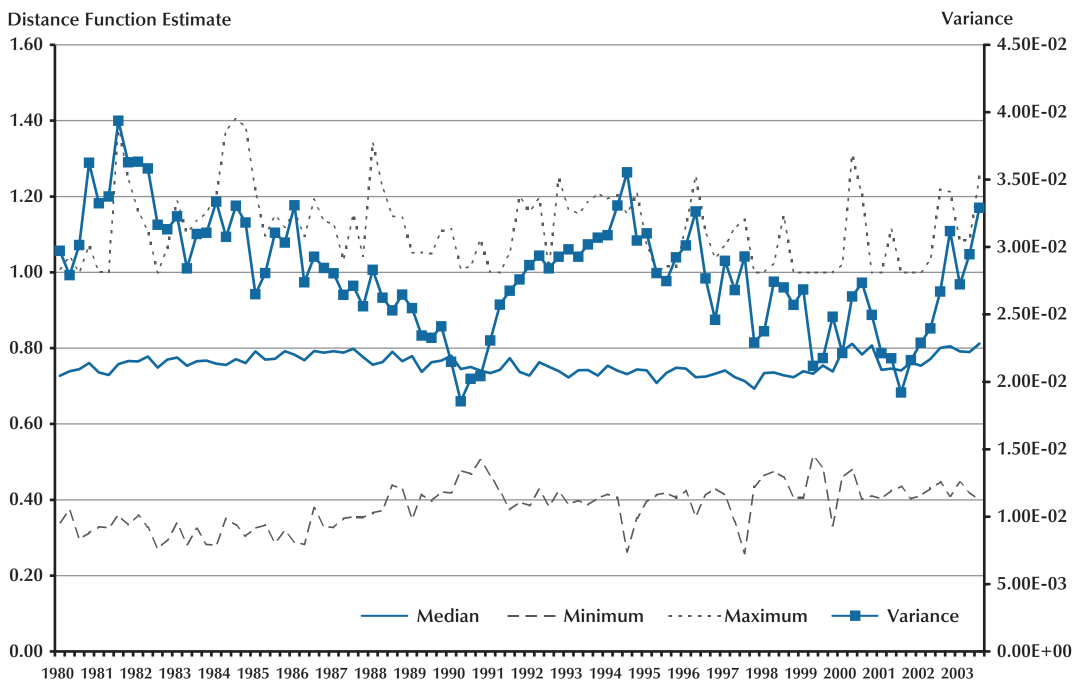


ously. A distance estimate larger than 1.0 indicates that the office uses more than the expected minimum, whereas a distance estimate less than 1.0 indicates that the office uses less than the expected minimum.

In general, the order-*m* inefficiency estimate for an office will be larger for larger values of *m*. Because the offices in the sample drawn to compute the order-*m* frontier produce at least as much output as the office in question, and because offices that produce more output typically use more input, the minimum input usage among *m* offices will tend to be smaller, the larger the number of offices sampled (i.e., the larger the value of *m*). Hence, the expected minimum input usage for any given office will generally be smaller, and the less efficient the office will appear to be, the larger the value of *m*. As Cazals, Florens, and Simar (2002, p. 7) note, the choice of *m* is arbitrary, but "a few values of *m* could be used to guide the manager of the production unit to evaluate its own performance" (p. 7). Here we are interested in the performance of Fed check offices in general, not specific offices, and the behavior of mean and median efficiency of check offices over time is

Figure 10

Order-*m* Efficiency, *m* = 5



largely invariant to the choice of m.[18] Moreover, as Figures 7 to 9 illustrate for 2003:Q3, efficiency estimates for individual offices using different values of m are highly correlated.

Although distance estimates and, hence, estimated order-m efficiency, vary with the value of m chosen, the trends in the mean and median values are similar across the four values of m—first increasing from 1980 to around 1987, then declining through about 1998, and finally rising through 2003. The changes in the trends are rather small, however, relative to the quarterly variability in the mean and median values, suggesting that the changes might not be statistically significant.

Consistent with the DEA efficiency estimates, Fed offices seem to have become somewhat less order-m efficient when the pricing regime was first introduced. For example, for $m = 5$, the median order-m distance estimate rises from approximately 0.73 in 1980 to 0.80 at its peak in 1987. In other words, in 1980, the input usage of the median office was 73 percent of the expected minimum amount, whereas in 1987, the input usage of the median office was 80 percent of the expected minimum amount. Order-m efficiency began to improve in the late 1980s, however, and continued to improve during much of the 1990s, a pattern consistent with the estimates of productivity obtained by Gilbert, Wheelock, and Wilson (2004). Mean and median estimates of order-m efficiency began to worsen around 1999, and also became somewhat more variable.

Mean and median efficiency estimates do not, of course, tell the whole story. Figure 10 plots the quarterly minimum, maximum, and median order-m efficiency estimates alongside the variance across Fed offices for $m = 5$ (plots for other values of m look similar). The plot reveals considerable variation over time in the minimum and, especially, maximum inefficiency estimates. In general, the variation in efficiency estimates across offices declined after the pricing regime was implemented in 1982, even

[18] Cazals, Florens, and Simar (2002) show that the order-m frontier converges to the FDH frontier as m approaches infinity. In our application, median order-m efficiency is very close to 1 in every period for $m = 50$, consistent with our finding that approximately 90 percent of all Fed offices are located on the FDH frontier in every period. The location of a high percentage of observations on the FDH (and DEA) frontier reflects a "curse of dimensionality" that plagues nonparametric frontier estimation when the combined number of inputs and outputs is high relative to the number of observations.

as mean and median inefficiency worsened through about 1987. After 1990, however, efficiency estimates became more dispersed until about 1994. Since then, dispersion has varied considerably from quarter to quarter.

DISCUSSION

We have examined the technical efficiency of Fed check offices over a long period that involved numerous environmental changes. The major environmental change of the early 1980s was the implementation of the pricing requirements of the Monetary Control Act of 1980. Both our DEA and order-m estimates suggest that, if anything, Fed offices became less efficient on average after full implementation of pricing in 1982. Efficiency had begun to improve by the late 1980s, however, and Fed offices generally became more efficient during the 1990s.

Federal Reserve check processing volume has been declining since 1999—a trend that is expected to continue, especially with enactment of the Check Clearing for the 21st Century Act. Our results indicate that the input-technical efficiency of Fed offices declined on average after 1999, with increased dispersion across offices. Because it is difficult and costly to reduce input amounts quickly, one would expect to observe a decrease in the efficiency of an office experiencing a sharp drop in check processing volume. Hence, one should be cautious about drawing strong conclusions about performance from short-run fluctuations in estimated efficiency.

While our results provide information about changes in the technical efficiency of the average (and median) Fed check office over time, it is important to note that we do not examine other types of efficiency or performance measures. For example, a measure of "overall" efficiency would capture both technical efficiency and allocative efficiency—which takes into consideration the feasibility of substituting among input types in response to shifts in relative input prices. One could also examine scale efficiency, which reflects the extent to which offices produce an efficient level of output.[19]

Further, while we have extended the standard model used to examine the efficiency of check office production by including the number of endpoints to which a check office delivers checks as a second type of output, our model cannot account for all differences in the operating environments or production of different check offices. Hence, as is the case with any empirical study of efficiency, differences in estimated efficiency across offices necessarily confound true differences in inefficiency that might be within the ability of managers to control with factors that are largely beyond control. For example, although we account for differences in the *number* of endpoints across offices, we do not control for differences in the geographic dispersion of endpoints served by Fed offices that could explain differences in their use of transportation input.[20] Nevertheless, while an examination of the causes of observed differences in estimated efficiency across offices is beyond the scope of the present study, the methodology described and illustrated in this article provides a framework for identifying differences that could help guide managers in their search for ways to control costs.

REFERENCES

Bauer, Paul W. and Ferrier, Gary D. "Scale Economies, Cost Efficiencies, and Technological Change in Federal Reserve Payments Processing." *Journal of Money, Credit, and Banking*, November 1996, *28*(4), pp. 1004-39.

Bauer, Paul W. and Hancock, Diana. "The Efficiency of the Federal Reserve in Providing Check Processing Services." *Journal of Banking and Finance*, April 1993, *17*(2-3), pp. 287-311.

Berger, Allen N. and Humphrey, David B. "Efficiency of Financial Institutions: International Survey and Directions for Future Research." *European Journal of Operations Research*, April 1997, *98*(2), pp. 175-212.

Board of Governors of the Federal Reserve System. *Annual Report*, 2000.

Board of Governors of the Federal Reserve System. *Annual Report*, 2002.

Cazals, Catherine; Florens, Jean-Pierre and Simar, Léopold. "Nonparametric Frontier Estimation: A Robust Approach." *Journal of Econometrics*, January 2002, *106*(1), pp. 1-25.

Gerdes, Geoffrey R. and Walton, Jack K. "The Use of Checks

[19] Gilbert, Wheelock, and Wilson (2004) examine scale efficiency and find no evidence that, as of 1999, any Fed office produced an inefficient level of output.

[20] Bauer and Hancock (1993) find that cost inefficiency of Fed check production was closely related to the geographic density of endpoints over their sample period.

and Other Noncash Payment Instruments in the United States." *Federal Reserve Bulletin*, August 2002, *88*(8), pp. 360-74.

Gilbert, R. Alton; Wheelock, David C. and Wilson, Paul W. "New Evidence on the Fed's Productivity in Providing Payments Services." *Journal of Banking and Finance*, September 2004, *28*(9), pp. 2175-90.

Kneip, Alois; Simar, Léopold and Wilson, Paul W. "Asymptotics for DEA Estimators in Non-parametric Frontier Models." Discussion Paper No. 0317, Institut de Statistique, Université Catholique de Louvain, Louvain-la-Neuve, Belgium, 2003.

McAllister, Patrick H. and McManus, Douglas A. "Resolving the Scale Efficiency Puzzle in Banking." *Journal of Banking and Finance*, April 1993, *17*(2-3), pp. 389-405.

Shephard, Ronald W. *Theory of Cost and Production Functions*. Princeton: Princeton University Press, 1970.

Simar, Léopold and Wilson, Paul W. "Statistical Inference in Nonparametric Frontier Models: The State of the Art." *Journal of Productivity Analysis*, January 2000, *13*(1), pp. 49-78.

Wheelock, David C. and Wilson, Paul W. "New Evidence on Returns to Scale and Product Mix Among U.S. Commercial Banks." *Journal of Monetary Economics*, June 2001, *47*(3), pp. 653-74.

Wheelock, David C. and Wilson, Paul W. "Robust Nonparametric Estimation of Efficiency and Technical Change in U.S. Commercial Banking." Working Paper No. 2003-037A, Federal Reserve Bank of St. Louis, November 2003.

Testing the Expectations Hypothesis: Some New Evidence for Japan

Daniel L. Thornton

I. INTRODUCTION

Spending decisions, especially investment decisions, are largely determined by long-term interest rates, while the actions of the monetary authority have a direct effect on interest rates only at the very short end of the yield curve. An important question in monetary economics and finance is, How do the actions of the monetary authority get translated along the entire yield curve? In countries where a wide variety of bonds with different maturities are traded, policy actions are thought to be translated from the short end to the long end of the term structure in accordance with the expectations hypothesis (EH), which asserts that the long-term rate is equal to market participants' expectation of the short-term rate over the holding period of the long-term asset plus a constant risk premium.

Until the mid-1980s, the Japanese bond market was relatively small, illiquid, and tightly regulated. Japan's capital markets were segmented by government regulations, not the public's preferences. Arbitrage opportunities across maturities were limited. Few thought that the EH applied to Japan. Consequently, there was no reason to test the EH for Japan. Monetary policy was thought to affect lending through quantity constraints and not through interest rates.

The Japanese began to deregulate their bond market in the 1970s. The structural changes in the Japanese financial markets have generated considerable interest in testing the EH using Japanese data (e.g., Campbell and Hamao, 1993; Singleton, 1990; Shirakawa, 1987; and Shikano, 1985). This paper is an extension of this research agenda, but differs from the previous literature in two respects. First, this paper investigates the EH by estimating a bivariate vector autoregression (VAR) for the long-term and short-term interest rates and testing the restrictions implied by the EH. This test was first suggested by Campbell and Shiller (1987); however, the procedure used here was developed by Bekaert and Hodrick (2001).[1]

Second, this paper deals directly with the issue of stationarity. While stationarity is frequently considered in testing the EH, its implications for the EH are seldom discussed. This paper attempts to fill this void. As a matter of theory, many economists and financial specialists appear to believe that interest rates are stationary. If they are not, the role played by the EH in monetary policy may be diminished.

As a practical matter, interest rates tend to exhibit considerable persistence. Indeed, the null hypothesis of nonstationarity is frequently not rejected even in relatively large, finite samples. This is particularly important for Japanese interest rates because they exhibit considerable persistence at the monthly frequency. It is well known, however, that tests of unit roots may have low power. Moreover, many financial market economists argue that interest rates are stationary on theoretical grounds. Given the differences of opinion about whether interest rates are stationary in general, the VAR test is applied under the assumption that interest rates are either stationary or nonstationary. The issue of stationarity is important only if the conclusions concerning the EH differ markedly depending on the assumption made.

[1] Several of the tests of the EH that are frequently used have low power and, more importantly, tend to generate results that can give a misleading impression of the strength or weakness of the EH. For a more detailed discussion of this problem, see Thornton (2002 and 2003a) and Kool and Thornton (2003).

Daniel L. Thornton is a vice president and economic advisor at the Federal Reserve Bank of St. Louis. This paper was prepared while the author was a visiting scholar at the Institute for Monetary and Economic Studies and was previously published (Institute for Monetary and Economic Studies, Bank of Japan, *Monetary and Economic Studies*, May 2004, *22*(2), pp. 45-69); it is reprinted here with permission. The views expressed are the author's and may not reflect the views of the Bank of Japan, the Federal Reserve Bank of St. Louis, or the Board of Governors of the Federal Reserve System. The author thanks Rob Dittmar, Hiroshi Fujiki, Marco Lippi, Takatoshi Ito, Shigenori Shiratsuka, Yi Wen, and an anonymous referee for helpful comments. Molly Dunn-Castelazo, John Zhu, and Kiyoshi Watanabe provided valuable research assistance.

Federal Reserve Bank of St. Louis *Review*, September/October 2004, *86*(5), pp. 21-39.

Japanese short-term interest rates, which began a rapid descent in late 1990, have hovered about their theoretical zero bound since mid- to late 1998. This feature of Japanese interest rates makes testing the EH after 1998 particularly difficult.

Section II discusses the evolution of Japanese financial markets during the postwar period and reviews the literature on tests of the EH in Japan in the postwar period. The data and some initial data analysis are presented in Section III. Section IV discusses nonstationarity and its implications for the EH. Some preliminary tests of the EH, which arise from the discussion in Section IV, are presented in Section V. In Section VI, the VAR test is applied to Japanese data under both assumptions—that interest rates are stationary or nonstationary, but cointegrated. The results of these tests are presented and discussed. The implications for rejecting the EH are presented in Section VII, and the conclusions are presented in Section VIII.

II. JAPANESE FINANCIAL MARKETS AND THE EH

A. *Evolution of Japanese Financial Markets*

The Japanese financial markets were highly regulated during the early postwar period. There was virtually no issuance of government debt during the first 15 years of the postwar period. Hence, there was little need for a government debt market. When, in the mid-1960s, the government needed to borrow to finance infrastructure, regulations were introduced that significantly limited the development of a secondary market in government debt. Specifically, banks were not permitted to resell government debt in the secondary market. Instead, there was an implicit guarantee that the Bank of Japan would purchase the government debt after a holding period of 1 year. In addition, securities companies were under administrative guidelines to maintain yields in the secondary market as close as possible to primary market yields. These restrictions significantly impeded the development of a secondary government debt market.[2]

Japanese corporations relied heavily on internal funds and loans from private financial institutions to finance investment. Equity and debt accounted for less than 5 percent of industrial funds prior to 1975 (Hodder, 1991).

The effect of the oil-price shock in the early 1970s facilitated the development of the *gensaki* market (the market for bond repurchase agreements) and the secondary market in government debt. A decrease in corporate investment following the oil-price shock led to an improvement in short-term corporate cash flows. This facilitated the expansion of the *gensaki* market, as firms sought alternatives to regulated bank deposits. The development of the *gensaki* market was enhanced further when the government formally recognized it and instituted prudential guidelines in 1976 (Takagi, 1988).

At the same time, deficit spending increased. The increased holding of government debt by banks prompted the Bank of Japan to suspend its commitment to repurchase government debt held for 1 year, which resulted in an erosion of bank liquidity. To shore up bank liquidity and to avoid debt monetization by the Bank of Japan, in April 1977 banks were permitted to sell government securities in the secondary market after a 1-year holding period. In addition, the requirement that secondary yields remain as close as possible to primary yields was lifted. Restrictions on banks' participation in the secondary market were eased further over time and eliminated in June 1985.

Simultaneous with these developments, the Japanese made a number of successive regulatory changes to liberalize cross-border capital flows. These and other steps toward financial deregulation resulted in deeper and more liquid financial markets, reduced transactions costs, and increased the substitution between assets with different characteristics and between assets with similar characteristics but different maturities (e.g., Singleton, 1990; Leung, Sanders, and Unal, 1991; and Takagi, 1988).

B. *The EH in Japan*

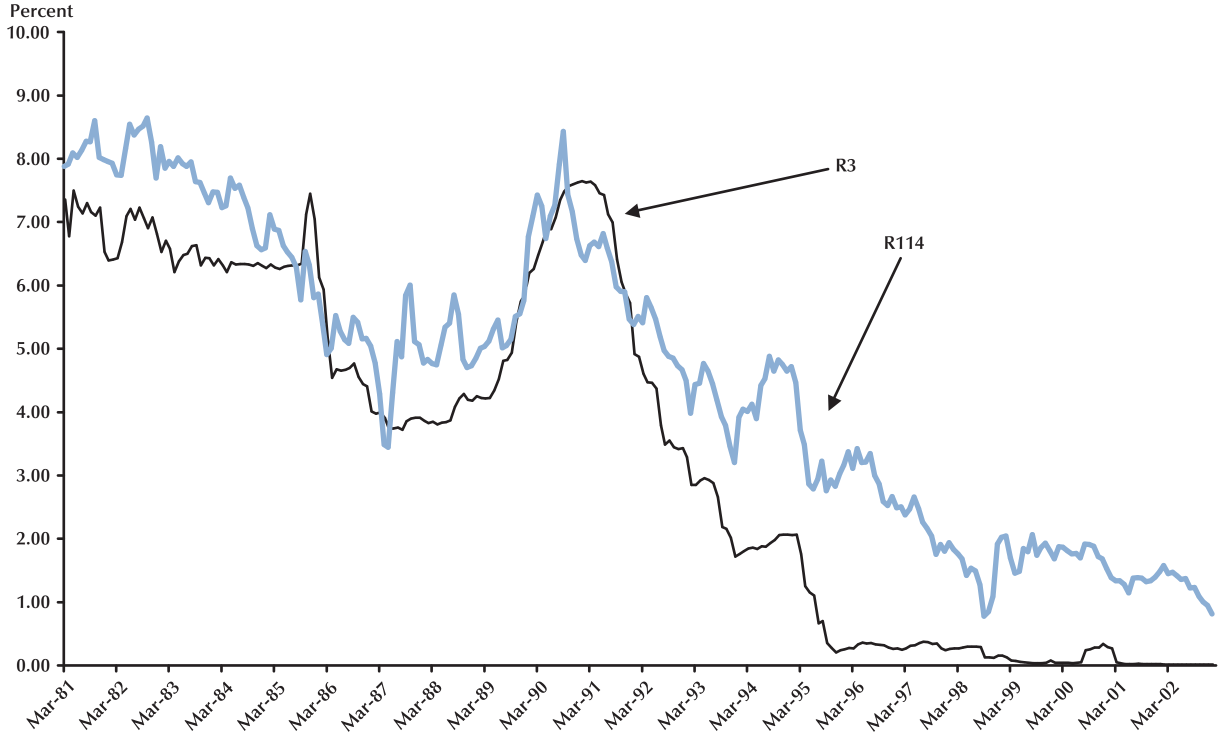
In financial markets where there is little possibility for arbitrage between assets with similar characteristics but of different maturities, there is little reason to think that the EH will hold. It is not surprising, then, that there was virtually no statistical testing of the EH using Japanese data prior to the mid-1980s. Financial market deregulation changed this. Since the 1980s, a number of researchers have tested the EH in Japan, where, as elsewhere, it is generally rejected.

Contrary to expectations, evidence supporting the EH appears to be strongest during periods when regulatory constraints effectively segmented long-term and short-term financial markets and weakest

[2] For additional details on these restrictions, see Takagi (1988).

Figure 1

3-Month *Gensaki* Rate and 114-Month Treasury Yield
(March 1981±January 2003)



when deregulation provided greater arbitrage opportunities. Shikano (1985) analyzes data from April 1977 to June 1984. He finds that the EH was not rejected for the entire period, but was rejected for the period October 1981 to June 1984. Similarly, Campbell and Hamao (1993) test the EH for the periods November 1980 through July 1985 and August 1985 through August 1990. While the EH is rejected for both periods, the qualitative evidence against the EH is stronger during the latter period. Similarly, Shirakawa (1987) finds that the EH fares worse for the period April 1981 through June 1986 compared with the period April 1977 through June 1986.

III. JAPANESE INTEREST RATES

A. Data

The data are end-of-month observations for the period March 1981 to January 2003. The rates include the 3-month *gensaki* rate and rates on Japanese government bonds (JGBs) with maturities of 0 to 1 year, 1 to 2 years, 2 to 3 years, up to 9 to 10 years. These data appear to be similar to those used by Campbell and Hamao (1993), which covered the period November 1980 to August 1990. Following Campbell and Hamao's taxonomy, the Treasury rates are designated as 6-, 18-, 30-, 42-, 60-, 78-, 90-, 102-, and 114-month rates, respectively. The *gensaki* rate was obtained from the Japan Securities Dealers Association, and the Treasury rates were obtained from the Bloomberg database. In cases where Treasury rates were missing, the Bloomberg data were supplemented with data compiled by the Bond Market Underwriter's Association.[3]

Figure 1 shows the 3-month *gensaki* rate and 114-month JGB yield over the March 1981–January 2003 period. Japanese rates declined generally until the late 1980s, rose until early 1990, and have since generally declined. Since the early 1990s, the *gensaki*

[3] I would like to thank Kiyoshi Watanabe for compiling these data. There was one missing observation for the 6-month rate that occurred on July 1992. The July observation was interpolated from the June and August 6-month rates.

rate has fallen much more rapidly than long-term rates. This is especially true during the first half of the 1990s, when the spread between the 114-month and 3-month rates increased dramatically. Moreover, since January 1992 the spread between the 114-month and 3-month rates has averaged nearly 175 basis points compared with about 60 basis points for the period up to January 1992.

B. The EH and the Zero Bound

The deterioration in the Japanese economy's performance in the early 1990s and the more recent deflation have greatly affected interest rates. Because of deflation, the *gensaki* rate has been at or near the zero nominal interest rate bound since late 1998. The zero bound has implications for testing the EH. Because market participants may still form expectations of the future behavior of the short-term interest rate (e.g., Okina and Shiratsuka, 2003), a zero interest rate policy may impact longer-term rates through the EH. What matters for the effectiveness of the EH is what Fujiki and Shiratsuka (2002) call the "policy duration effect"—i.e., how long the market anticipates that the monetary authority will maintain the current target rate—not whether monetary policy actions are anticipated (e.g., Thornton, 2003b). It is nevertheless the case that, when the short-term rate is at the zero bound, the spread between the long-term and short-term rate need not provide information about the direction of changes in the short-term real rate.

C. Persistence of Japanese Interest Rates

Like U.S. interest rates, Japanese interest rates exhibit considerable persistence. The results of augmented Dickey-Fuller tests (Dickey and Fuller, 1979) are reported in Table 1. Because the qualitative conclusions were sensitive to the choice of lag length, the lag lengths were chosen by the Schwarz criterion and are denoted in parentheses below the Dickey-Fuller test statistic. The results indicate that the null hypothesis of a unit root is not rejected even at the 10 percent significance level for any of the rates.

The conclusions are robust over the sample period. This is illustrated in Figure 2, which shows the results from applying the augmented Dickey-Fuller test to a rolling sample of 78 monthly observations for the 3-, 66-, and 114-month rates. The results for the other Treasury rates are similar to those shown in Figure 2. In general, the degree of persistence increases with the maturity of the rate. Except for some relatively short periods, and primarily for the *gensaki* and other shorter-term rates, the null hypothesis of a unit root is not rejected. It is well known that the augmented Dickey-Fuller test is sensitive to shifts or breaks in the time-series process. Hence, the conclusion that rates have a unit root is not too surprising given the marked decline in the level of rates in mid-1996. However, the unit root hypothesis is also not rejected for any of the rates for the seemingly more stable period from March 1981 to October 1990.

D. Volume of Trade

It is important to note that, like the U.S. market for Treasury securities, trading in Japanese Treasury securities is focused on maturities at or near that of the benchmark issue. When the trading volume on a particular issue of Treasury debt with approximately 10 years of term remaining becomes large enough, it might be designated by the market to be the "benchmark" issue. An issue remains the benchmark issue until another issue receives this designation. There have been a number of benchmark issues since August 1983. Higo (2000) reports that the average maturity of an issue when it was first designated the benchmark issue is just under 10 years. Trading volume in the Japanese Treasury market tends to be focused on the benchmark issue, maturities close to that of the benchmark issue, and, to a much lesser extent, previous benchmark issues regardless of their remaining maturity (Higo, 2000). The market for other issues is relatively thin.

If the markets are thin, there may be day-to-day or month-to-month variation in the rate that is due solely to random variation owing to the thinness of the market. One way to investigate whether the thinness of the markets for shorter-maturity assets might impact our results is to determine whether the variance of the rates declines as the maturity approaches that of the benchmark issue. This is done here by calculating the ratio of the variance of the rate at each maturity to the variance of the 3-month rate. If the thinness of the market is an important factor, one might expect to see a marked drop in the variance ratio as the maturity approaches 10 years.

Because of uncertainty as to whether interest rates are stationary or unit root processes, the variance ratios are calculated for the levels and first differences of the rates. The variance ratios for levels and first differences are shown in Figures 3 and 4,

Table 1

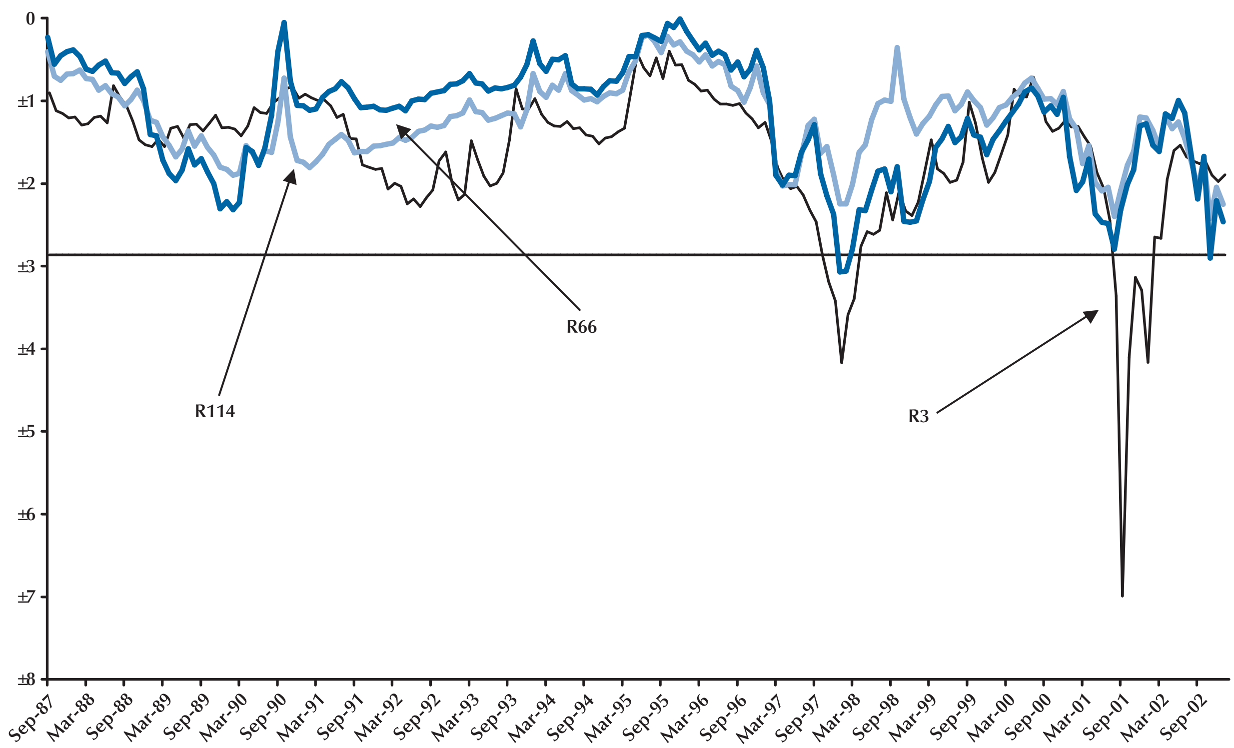
Augmented Dickey-Fuller Tests for Japanese Treasury Rates

	R3	R6	R18	R30	R42	R54	R66	R78	R90	R102	R114
March 1981–January 2003											
DF	–1.13	–1.00	–1.51	–1.14	–1.09	–0.96	–0.88	–1.32	–1.28	–0.97	–0.86
	(3)	(0)	(2)	(0)	(0)	(0)	(0)	(1)	(0)	(0)	(0)
March 1981–October 1990											
DF	–1.28	–1.23	–1.33	–1.38	–1.31	–1.22	–1.11	–1.30	–1.56	–1.44	–1.50
	(4)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)

NOTE: Parentheses indicate lag length.

Figure 2

Rolling Augmented Dickey-Fuller Test for Selected Rates

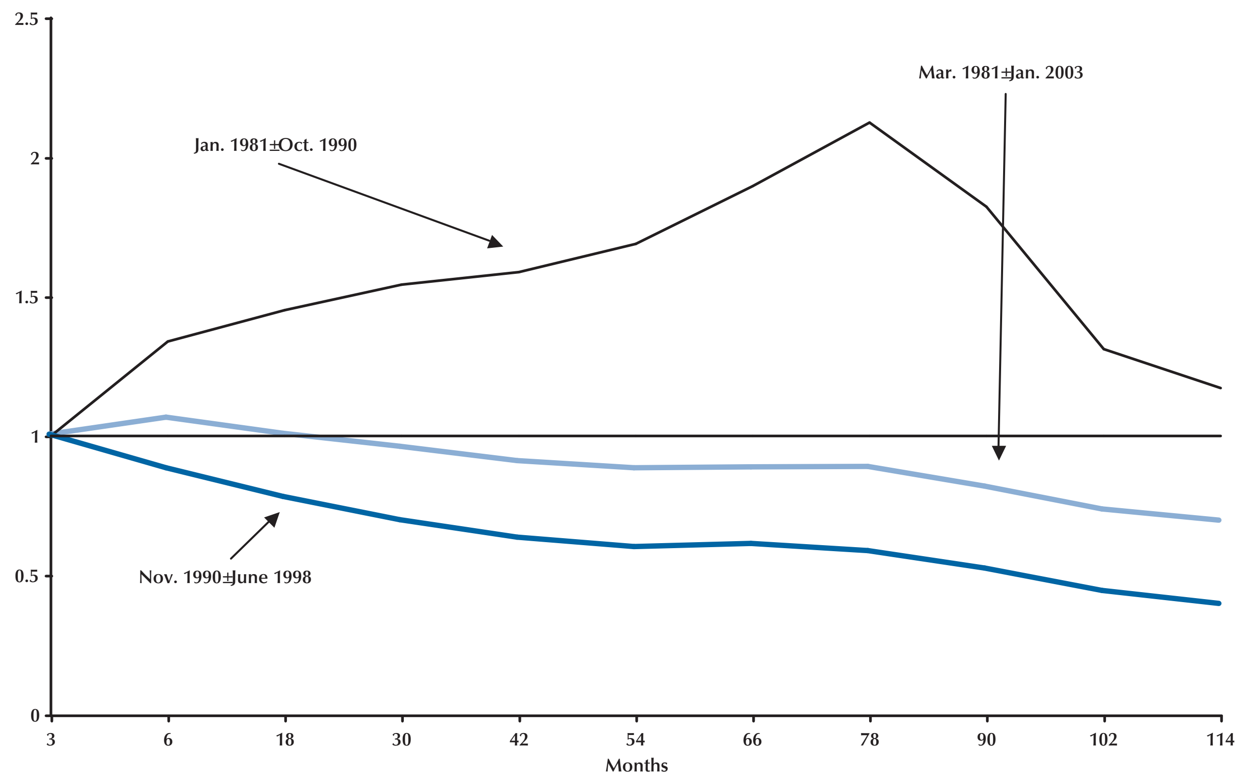


respectively, for selected periods. Figure 3 suggests the possibility of a thin-market effect for the period March 1981–October 1990, when the variance ratio increases to a maturity of 78 months and then declines markedly. The variance ratios for the entire period and for the period since October 1990 provide no indication of a thin-market effect. The decline of the variance ratios is nearly monotonic, as one would expect if there were an important risk premium. The curve for the entire sample lies above that for the period ending in 1998, because after short-term rates achieved the zero bound, long-term rates became much more variable than short-term rates.

The variance ratios for the first-difference data also give no indication of a significant thin-market

Figure 3

Ratio of the Variance of Each Rate to the Variance of R3 for Selected Periods



effect. While the variance ratio first rises and then declines during the November 1990–June 1998 period, the sharp decline in the variance occurs before the maturity approaches 10 years. This simple analysis suggests that the thinness of the market may not be important for this research. If it is, it would appear to be important only for the level of rates and for the period March 1981–October 1990.

IV. IMPLICATIONS OF NONSTATIONARITY FOR THE EH

A. Nonstationarity

Given the evidence of nonstationarity in Japanese interest rates, it is important to ask, What, if anything, does nonstationarity imply for the EH? To answer this question, assume that the stochastic process driving the short-term rate is

$$r_t^m = \rho r_{t-1}^m + \varepsilon_t , \tag{1}$$

where r_t^m is the current value of the short-term, m-period rate; ρ is a parameter such that $0 \le \rho \le 1$; and ε_t is an *i.i.d.* random variable distributed with a mean zero and a variance σ^2. If $0 \le \rho < 1$, the short-term interest is generated by a stationary stochastic process. If $\rho = 1$, on the other hand, the stochastic process is said to be I(1). In this case, stationarity is achieved by first-differencing the short-term rate.

The EH assumes that

$$r_t^n = (1/k)\sum_{i=0}^{k-1} E_t r_{t+mi}^m + \pi , \tag{2}$$

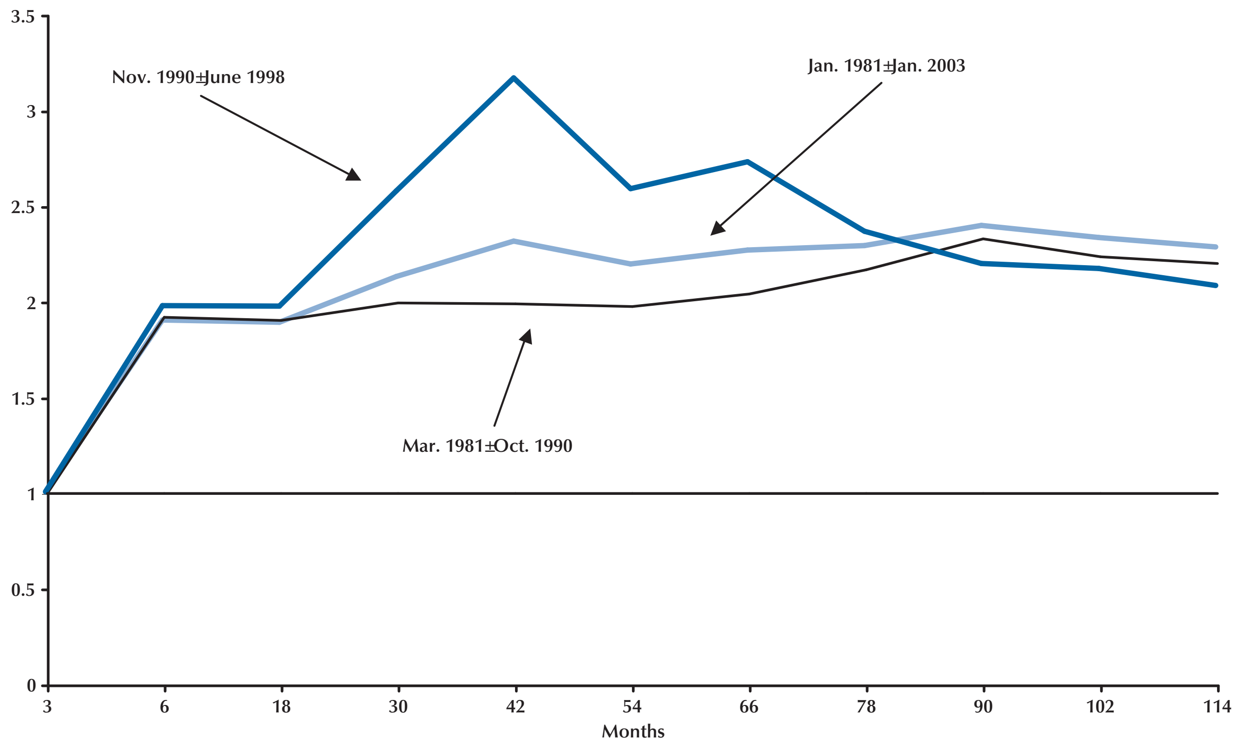
where r_t^n is the long-term, n-period rate, and $k = n/m$ is an integer. If the short-term rate is a unit root process, $E_t r_{t+mi}^m = r_t^m$ for all i. Note that $E_t r_{t+mi}^m$ equals r_t^m rather than r_{t-1}^m, because in the EH literature it is assumed that the short-term rate is observed when the long-term rate is determined. Substituting into the above expression and simplifying yields

$$r_t^n = r_t^m + \pi . \tag{3}$$

Therefore, if the short-term rate is I(1) and the EH holds, the long-term rate would always equal the

Figure 4

Ratio of the Variance of the First Difference of Each Rate to the Variance of the First Difference of R3 for Selected Periods



short-term rate plus a constant risk premium. This result stems from the fact that, if the short-term rate is I(1), the best estimate of the short-term rate at any horizon is its current level—changes in the short-term rate are unpredictable, i.e., $\Delta r_{t+i}^{m} = \varepsilon_{t+i}$ for all $i \geq 1$.[4]

The EH is useful to market participants and policymakers because, if it holds, the spread between the long-term rate and the short-term rate provides information about the future level of the short-term rate. If the short-term rate is impossible to predict, however, the spread between long-term and short-term rates cannot provide useful information about the market's expectation for the short-term rate. In this case, the EH is of little practical use, even though (as in the example above) it holds.

B. Cointegration

The above analysis ignores the possibility that the short-term and long-term rates are cointegrated. If these rates are unit root processes, but cointegrated, they are stationary in levels in the direction of the cointegrating vector. The idea of cointegration is illustrated by assuming that the long-term and short-term interest rates are jointly endogenous and that the true data-generating process can be approximated by a VAR of the form

(4) $$y_t = \Theta(L) y_{t-1} + \eta_t ,$$

where $y_t = (r_t^m, r_t^n)$ and $\Theta(L)$ is a P-order polynomial in the lag operator L. Equation 4 can be written as

(5) $$\Delta y_t = \Psi(L) \Delta y_{t-1} - \Pi y_{t-1} + \eta_t ,$$

where $\Pi = (I - \Theta(1))$. If y_t is stationary, the rank of Π is 2. In this case, any linear combination of the long- and short-term rates is stationary. If the short- and long-term rates have a unit root, however, the rank

[4] If the short-term rate is not generated by a simple I(1) process (as in this example) but is nonstationary, the general idea still applies because the variance of the short-term rate is not finite. It would be the case, however, that there would be some predictability of changes in the short-term rate. Generally speaking, the degree of predictability will be positively related to the extent to which the root is greater than 1.

of Π is at most 1. If the rank of Π is 1, the long- and short-term rates are cointegrated. In this case, $\alpha\beta' = \Pi$, where α and β are 2×1 vectors. The cointegrating vector, β, represents the long-run equilibrium relationship between the long- and short-term interest rates, i.e., the direction in which the relationship between the levels of the rates is stationary. Specifically, $\beta' y_t$ is stationary (mean reverting). Hence, cointegration indicates that there is a stable equilibrium relationship between the levels of the short- and long-term interest rates, but only in the direction of the cointegrating vector. When the variables are cointegrated, $\alpha\beta' y_{t-1}$ replaces Πy_{t-1} in (5) and the resulting equation is referred to as an error-correction model (ECM), where the coefficients in α measure the speed with which the rates adjust to their long-run equilibrium.

C. Cointegration and the EH

If the rates are nonstationary but cointegrated, one can test the EH by testing the hypothesis that the cointegrating vector (adjusted for the constant risk premium and/or a deterministic trend) equals (1, –1). This test has been used by a number of researchers (e.g., Stock and Watson, 1988; Hall, Anderson, and Granger, 1992; Engsted and Tanggaard, 1994; and Sarno and Thornton, 2003).

Lack of cointegration is relatively strong evidence against the EH for two reasons. First, if the interest rates are truly I(1), rejecting the hypothesis of cointegration implies that there is no stable long-run relationship between the levels of the interest rates, that is, the EH cannot hold.

Second, it is well known that the power to reject the null hypothesis of nonstationarity is low when the root is close to 1. Thus, it could be that interest rates are really I(0). If this is the case, however, it should not be too difficult to find evidence of cointegration, i.e., reject the null hypothesis that there is no stationary relationship between the long-term rate and the short-term rate. Therefore, failure to find evidence of at least one cointegrating relationship among stationary variables is relatively strong evidence against the EH.

Finding cointegration, but rejecting the hypothesis that the cointegrating vector is (1, –1), is also relatively strong evidence against the EH because it suggests that the equilibrium relationship is in a direction that is inconsistent with the EH. If the EH does not hold in the long-run equilibrium, there is little reason to expect that it will hold at frequencies that are of interest to policymakers and financial analysts.

On the other hand, finding that the EH holds in the long run does not necessarily imply that the EH is useful for policymakers and market analysts. To be useful, longer-term rates must respond reasonably quickly to changes in the policy rate. Hence, failing to reject the hypothesis that the cointegrating vector is (1, –1) does not establish that the EH holds at frequencies that are of interest to policymakers and financial market analysts. Policymakers and financial analysts need to know how quickly the long-term rate can be expected to adjust to policy-induced changes in the short-term rate.

V. TESTING THE EH FOR JAPANESE TREASURY RATES

If the EH holds, on average, the long-term rate will equal the short-term rate plus a constant risk premium. A simple way to test the EH is to test for a unit root in the spread between the long-term and short-term rates.[5] Figure 5 presents spreads between 6-month and 114-month Treasury rates and the *gensaki* rate. Other spreads tend to lie within the boundaries established by these spreads. Both spreads exhibit considerable persistence, suggesting the possibility that the null hypothesis of nonstationarity will not be rejected.

Augmented Dickey-Fuller tests of all of the spreads, for the entire sample period and for the March 1981–October 1990 period, are presented in Table 2. As before, the lag lengths were chosen by the Schwarz criterion and are denoted in parentheses below the Dickey-Fuller test statistic. The unit root hypothesis is rejected for all rate spreads except the 114-month rate for the entire sample period and for some of the intermediate maturities for the period ending October 1990.

The temporal stability of this conclusion is investigated by a rolling simple Dickey-Fuller test for the rate spreads using a sample of 78 monthly observations. The results for R114, R54, and R6 are presented in Figure 6. R54 is shown because the test statistics for rates with maturities longer than 42 months were generally larger than those for the 42-month rate. The results suggest that the null hypothesis of a unit root is somewhat borderline for the spread between the 42-month Treasury rate

[5] See Dickey, Jansen, and Thornton (1991, pp. 59-60) for a discussion of why this is appropriate.

Figure 5

Selected Rate Spreads

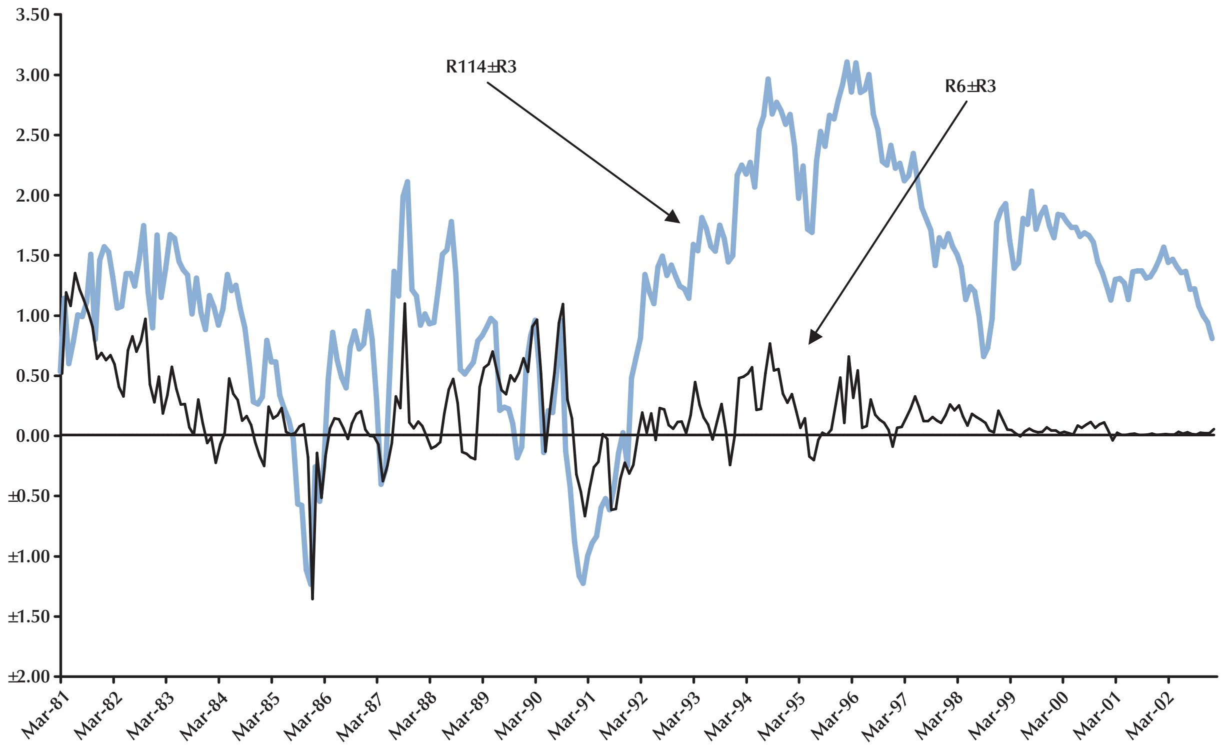


Table 2

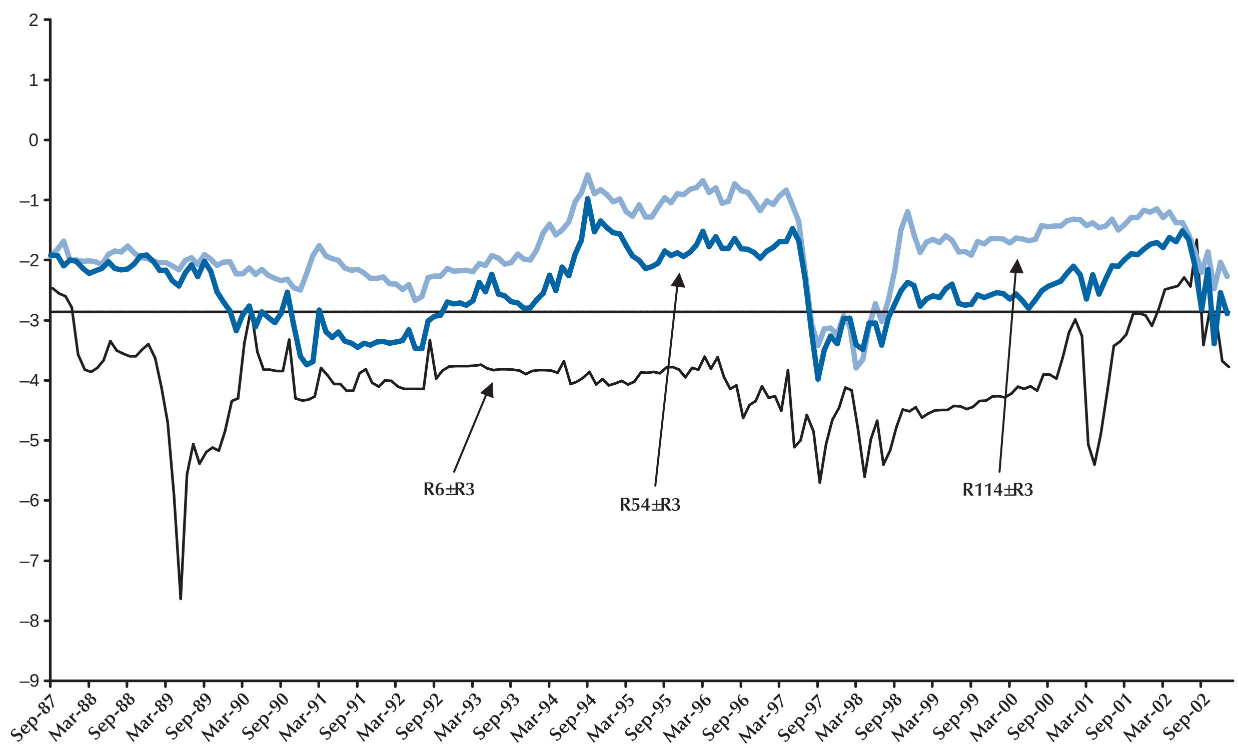
Augmented Dickey-Fuller Tests of Japanese Treasury Rate Spreads

Long rate	R6	R18	R30	R42	R54	R66	R78	R90	R102	R114
March 1981–January 2003										
DF	–5.50*	–4.94*	–4.47*	–4.07*	–3.48*	–3.15*	–3.15*	–3.40*	–2.95*	–2.79
	(1)	(1)	(1)	(0)	(0)	(0)	(1)	(1)	(0)	(0)
March 1981–October 1990										
DF	–3.66*	–3.46*	–3.35*	–3.25*	–2.84	–2.47	–2.65	–3.51*	–3.42*	–3.17*
	(1)	(1)	(0)	(0)	(0)	(0)	(1)	(1)	(0)	(0)

NOTE: Parentheses indicate lag length; * indicates significance at the 5 percent level.

Figure 6

Rolling Dickey-Fuller Test for Selected Interest Rate Spreads



and the *gensaki* rate, but is infrequently rejected for Treasury rates with maturities longer than 42 months and frequently rejected for rates with maturities of less than 42 months. Furthermore, it appears that the null hypothesis is nearly always rejected when the long-term rate is 6 months. These results suggest that the EH may not hold when the long-term rate is 42 months or longer.

This conclusion is supported by formal tests for cointegration for three periods—the entire sample period, the period ending October 1990, and the period from November 1990–June 1998, before short-term rates reached the zero bound.[6] The lag order and the precise form of the cointegration model are jointly determined by the Schwarz criterion. These results suggest that the 6-, 18-, 30-, and 42-month rates are cointegrated with the 3-month *gensaki* rate. The cointegration test results for rates with maturities of longer than 42 months indicate that these rates are not cointegrated with the *gensaki* rate for either the entire sample period or for the November 1990–June 1998 period. The results for the March 1981–October 1990 period are mixed. There are model specifications for which the null hypothesis of no cointegrating vector is rejected. These are not the specifications that minimized the Schwarz criterion, however. Moreover, when these models are estimated, the restriction that the cointegrating vector is (1, –1) is easily rejected. For these reasons, the cointegration test results for maturities longer than 42 months are not presented.

Estimates of the cointegrating vectors for rates up to 42 months are presented in Table 3. The coefficient estimates are normalized on the short-term rate, so Table 3 reports the estimate of the coefficient on the long-term rate and the χ^2 test statistics for the null hypothesis that the cointegrating vector is (1, –1). The estimated coefficients are close to –1 for the entire sample period and for the

6 Several alternative specifications of the cointegration model (allowing for a constant term and/or a deterministic trend in the cointegration relationship and no trend, a linear trend, or a quadratic trend in the structural dynamics) and alternative lag lengths from one through three were considered. The null hypothesis of at least one cointegrating vector is rejected in every case for the entire sample period and for the March 1981–October 1990 subperiod.

Table 3

Estimated Cointegrating Vectors

	R6	R18	R30	R42
March 1981–January 2003				
Coefficient	–0.962	–1.021	–0.973	–0.961
	(0.01)	(0.02)	(0.03)	(0.05)
χ^2	8.155	0.715	0.500	0.370
	[0.004]	[0.399]	[0.479]	[0.543]
March 1981–October 1990				
Coefficient	–0.96	–0.957	–0.960	–0.952
	(0.01)	(0.02)	(0.02)	(0.02)
χ^2	6.077	4.079	2.426	2.747
	[0.014]	[0.043]	[0.119]	[0.097]
November 1990–June 1998				
Coefficient	–1.171	–1.308	–1.488	–1.784
	(0.04)	(0.07)	(0.12)	(0.19)
χ^2	14.476	16.751	14.082	14.903
	[0.000]	[0.000]	[0.000]	[0.000]

NOTE: Parentheses indicate standard errors; brackets indicate significance levels.

period March 1981–October 1990. The coefficient is more precisely estimated for the 6- and 18-month rates, so the null hypothesis that the cointegrating vector is (1, –1) is rejected for the 6-month rate for the entire sample period and for both the 6- and 18-month rates for the period March 1981–October 1990. In these cases, however, the departure of the equilibrium relationship from that which is consistent with the EH is not large. At the shorter end of the maturity spectrum over the entire sample period and for the March 1981–October 1990 period, the equilibrium relationships appear to be more or less consistent with the EH holding in the long run.

The EH is easily rejected for the November 1990–June 1998 period, however. The point estimates of the coefficients on the long-term rate are very far from –1. Moreover, the null hypothesis that the cointegrating vector is (1, –1) is rejected for all rates at very low significance levels. Hence, as with previous empirical work, there is no evidence that long-term rates behave in a manner consistent with the EH during the more recent sample period.

VI. VAR TEST OF THE EH

Campbell and Shiller (1987) suggest that the EH be tested by testing the restrictions imposed by the EH on a VAR of the short-term and long-term interest rates. The restrictions implied by the EH are highly nonlinear, however, and the Wald test, which they used, is known to be affected greatly by nonlinearity. Consequently, they suggested that the major advantage of their VAR approach came from its ability to generate economic measures of the relative importance of the EH.

Bekaert and Hodrick (2001) propose a method for testing the restrictions imposed on a VAR by the EH using a Lagrange multiplier (LM) test. Since this procedure is relatively new, it is outlined here in some detail. The test is illustrated using a VAR expressed in levels; however, only minor changes are required to use the Bekaert-Hodrick procedure to test Campbell and Shiller's (1987) specification.

A. Bekaert and Hodrick Test

This test is general and can be applied to any VAR specification where the restrictions implied by the EH can be imposed. To illustrate the procedure, it is assumed that interest rates are stationary, so that the VAR takes the form of (4), i.e.,

$$(I - \Theta(L))y_t = \eta_t \tag{6}$$

for $y_t = (r_t^m, r_t^n)'$. Generalized method of moments (GMM) estimation imposes orthogonality conditions of the form $g(z_t,\theta) \equiv \eta_t \otimes x_{t-1}$, where x_{t-1} is a vector formed from stacking lagged values of y_t, possibly with a constant, z_t is defined as $(y_t, x_{t-1})'$, and θ is a vector formed from the parameters in $\Theta(L)$. Using the sample moment condition,

$$(7) \qquad g_T(\theta) \equiv \frac{1}{T}\sum_{t=1}^{T} g(z_t,\theta),$$

GMM estimation proceeds by choosing θ to minimize the following objective:

$$(8) \qquad J_T(\theta) \equiv g_T(\theta)' W g_T(\theta).$$

The optimal weighting matrix, W, is a consistent estimator of the inverse of

$$(9) \qquad \Omega \equiv \sum_{k=-\infty}^{k=\infty} E\left[g(z_t,\theta)g(z_{t-k},\theta)'\right].$$

GMM is used to estimate restricted VARs by forming a Lagrangian from the usual GMM quadratic objective and a vector of parameter constraints. The Lagrangian is defined

$$(10) \qquad L(\theta,\gamma) = -\frac{1}{2} g_T(\theta)'\Omega_T^{-1} g_T(\theta - a_T(\theta)'\gamma,$$

where γ is a vector of Lagrange multipliers and the constraints on θ have been represented by the vector-valued function, $a_T(\theta) = 0$. Here the matrix Ω_T is again a consistent estimate of the matrix Ω defined above. Denoting the Jacobian of $g_T(\theta)$ and $a_T(\theta)$ by G_T and A_T, respectively, the first-order conditions for maximizing $\bar{\theta}$ and $\bar{\gamma}$ can be written as

$$(11) \qquad \begin{bmatrix} -G_T{}'\Omega_T^{-1}\sqrt{T}g_T(\bar{\theta}) - A_T{}'\sqrt{T}\bar{\gamma} \\ -\sqrt{T}a_T(\bar{\theta}) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

The asymptotic distribution of the constrained estimator can be derived from these first-order conditions by expanding $g_T(\theta)$ and $a_T(\theta)$ in Taylor series around the true parameter value, θ_0, and substituting these into the first-order conditions above. This yields a system of the form

$$(12) \qquad \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -G_T{}'\Omega_T^{-1}\sqrt{T}g_T(\theta_0) \\ 0 \end{bmatrix} - \begin{bmatrix} B_T & A_T{}' \\ A_T & 0 \end{bmatrix} \begin{bmatrix} \sqrt{T}\left(\bar{\theta} - \theta_0\right) \\ \sqrt{T}\bar{\gamma} \end{bmatrix},$$

for $B_T \equiv G_T{}'\Omega_T^{-1}G_T$. Use of the partitioned inverse formula allows one to argue that the constrained estimator, $\bar{\theta}$, is distributed as $\sqrt{T}(\bar{\theta} - \theta_0) \to N(0, \Sigma_T)$ for

$$(13) \qquad \Sigma_T \equiv B_T^{-1} - B_T^{-1}A_T{}'\left(A_T B_T^{-1} A_T{}'\right)^{-1} A_T B_T^{-1}$$

and the Lagrange multipliers are distributed asymptotically as

$$(14) \qquad \sqrt{T}\bar{\gamma} \to N\left(0, \left(A_T B_T^{-1} A_T{}'\right)^{-1}\right).$$

If the constraints have a significant impact on parameter estimation, then the estimated Lagrange multipliers should be significantly different from zero. The asymptotic distributions given above can be used to show that a test that the multipliers are jointly zero can be based on the statistic

$$(15) \qquad T\bar{\gamma}'\left(A_T B_T^{-1} A_T{}'\right)\bar{\gamma},$$

which is asymptotically distributed as $\chi^2(l)$, where l is the number of restrictions imposed.

Maximization of the Lagrangian above is often computationally troublesome, so Taylor series approximations to $a_T(\theta)$ and $g_T(\theta)$ can again be used to derive a constrained estimate with similar asymptotic properties. Instead of expanding around the true value, θ_0, the current estimate of the true value, θ_i, is used to form a better approximation, θ_{i+1}. Since

$$g_T(\theta_{i+1}) \approx g_T(\theta_i) + G_T(\theta_{i+1} - \theta_i)$$

and

$$a_T(\theta_{i+1}) \approx a_T(\theta_i) + A_T(\theta_{i+1} - \theta_i),$$

we can substitute in the first-order conditions for maximization to derive the following iterative method:

$$(16) \qquad \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -G_T{}'\Omega_T^{-1}\sqrt{T}g_T(\theta_i) \\ -\sqrt{T}a_T(\theta_i) \end{bmatrix} - \begin{bmatrix} B_T & A_T{}' \\ A_T & 0 \end{bmatrix} \begin{bmatrix} \sqrt{T}\left(\theta_{i+1} - \theta_i\right) \\ \sqrt{T}\gamma_{i+1} \end{bmatrix}.$$

The unconstrained VAR parameter estimates are used for the initial conditions and the procedure iterates until the constraints are satisfied. The moment conditions for VAR estimation should be uncorrelated over time. Hence, Ω_T is estimated by

$$(17) \qquad \Omega_T = \frac{1}{T}\sum_{t=1}^{T} g(z_t,\theta_U) g(z_t,\theta_U)',$$

evaluating the moment conditions at the unconstrained VAR parameter estimates.

The constraints that the EH imposes on a VAR can be seen by writing the VAR in first-order form, that is,

$$(18)\quad \begin{pmatrix} r_t^m \\ r_t^n \\ r_{t-1}^m \\ r_{t-1}^n \\ \vdots \\ r_{t-k}^m \\ r_{t-k}^n \end{pmatrix} = \begin{pmatrix} \theta_1 & \theta_2 & \cdots & & \theta_r \\ I & 0 & & & 0 \\ 0 & I & & & \vdots \\ \vdots & & \ddots & & \\ 0 & & & I & 0 \end{pmatrix} \begin{pmatrix} r_{t-1}^m \\ r_{t-1}^n \\ r_{t-2}^m \\ r_{t-2}^n \\ \vdots \\ r_{t-k-1}^m \\ r_{t-k-1}^n \end{pmatrix} + \begin{pmatrix} \eta_t \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

or simply $x_t = \Theta x_{t-1} + \upsilon_t$. Note that $E_t(x_{t+k}) = \Theta^k x_t$, so that $E_t(r_{t+k}^m) = e_1{}'\Theta^k x_t$ for $e_1 = (1,0,\ldots,0)'$. Note, too, that $r_t^n = e_2{}'x_t$ for $e_2 = (0,1,0,\ldots,0)'$. Consequently, for any two interest rates such that $k = n/m$ is an integer, the EH implies that

$$(19)\qquad r_t^n = \frac{1}{k}\sum_{i=0}^{k-1} E_t\left(r_{t+mi}^m\right),$$

so that the EH can be expressed equivalently as

$$(20)\qquad e_2{}'x_t = \frac{1}{k}\sum_{i=0}^{k-1} e_1{}'\Theta^{mi} x_t.$$

The constraints that satisfy the EH are given by

$$(21)\qquad a_T(\theta) \equiv e_2{}' - \frac{1}{k}\sum_{i=0}^{k-1} e_1{}'\Theta^{mi} = 0.$$

No simple closed form exists for the Jacobian of these constraints. Consequently, they are calculated numerically for use in the iterative procedure described above.

B. Campbell and Shiller Test and Cointegration

Campbell and Shiller's (1987) specific proposal stems from a concern that interest rates are nonstationary. Their test is based on the fact that (2) can be rewritten as

$$(22)\qquad S_t = E_t \sum_{i=1}^{k-1} (1 - i/k)\Delta^m r_{t+mi}^m,$$

where $S_t = (r_t^n - r_t^m)$ and Δ^m denotes the m-horizon change, i.e., $\Delta^m w_t = w_{t+m} - w_t$. Specifically, Campbell and Shiller propose estimating a VAR representation,

$$(23)\qquad x_t = A(L)x_{t-1} + \omega_t,$$

where $x_t = (\Delta r_t^m, S_t)'$ and $A(L)$ is a P-order polynomial in the lag operator L, and testing the restrictions implied by equation 22. Noting that (23) can be rewritten as

$$(24)\quad \begin{pmatrix} x_t \\ x_{t-1} \\ \vdots \\ x_{t-P} \end{pmatrix} = \begin{pmatrix} A_1 & A_2 & \cdots & & A_P \\ I & 0 & & & 0 \\ 0 & I & & & \vdots \\ \vdots & & \ddots & & \\ 0 & & & I & 0 \end{pmatrix} \begin{pmatrix} x_{t-1} \\ x_{2-1} \\ \vdots \\ x_{t-P-1} \end{pmatrix} + \begin{pmatrix} \omega_t \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

or more compactly as

$$(25)\qquad x_t^* = Ax_{t-1}^* + \omega_t,$$

Campell and Shiller (1987, 1991) note that (22) can be written as

$$(26)\quad S_t = e_1'A[I - (m/n)(I - A^n)(I - A^m)^{-1}](I - A)^{-1} x_t^*.$$

Hence, the EH can be tested under the assumption that interest rates are nonstationary by testing the restriction

$$(27)\quad e_2' - e_1'A[I - (m/n)(I - A^n)(I - A^m)^{-1}](I - A)^{-1} = 0.$$

It should be noted that (4) and (23) are comparable ($\omega_t = \eta_t$) if and only if the cointegrating vector is $(1,-1)$. In this case, Campbell and Shiller's test preserves the level relationship between the long-term and short-term rates because, under these conditions, (23) can be derived from simple algebraic manipulations of (4).[7] Thus, in situations where the long-term and short-term rates satisfy the necessary conditions for the EH holding in the long run, Campbell and Shiller's specification provides a way of testing whether the EH holds at frequencies that are of interest to policymakers.[8]

C. Results of VAR Tests

The Bekaert-Hodrick procedure is applied to the VAR of the form of (4) under the assumption that interest rates are stationary. Because the LM test using the level specification is valid only if the VAR is stable, the maximum eigenvalue for each of the unrestricted VARs is calculated. In all instances, the maximum eigenvalue is less than 1. The VAR test is also applied to the Campbell-Shiller VAR, (23), under the assumption that interest rates are non-

[7] See Thornton (1985, especially the appendix) and Chow (1964) for a discussion of the role of normalization in regression analyses.

[8] This test has been employed using a Wald test by Campbell and Shiller (1987) and Carriero, Favero, and Kaminska (2003) and using the LM test by Bekaert and Hodrick (2001) and Dittmar and Thornton (2003).

Table 4

LM Statistics for LM Test Using Level Data

	March 1981–January 2003		March 1981–October 1990		November 1990–June 1998	
	R3	R6	R3	R6	R3	R6
R6	**139.228**	—	**119.619**	—	**68.705**	—
	(0.000) 2		(0.000) 1		(0.000) 1	
R18	**14.315**	5.903	**9.562**	5.467	**8.257**	**10.548**
	(0.006) 2	(0.052) 1	(0.008) 1	(0.065) 1	(0.016) 1	(0.005) 1
R30	**11.409**	**13.613**	**6.402**	**14.601**	**8.190**	**11.539**
	(0.022) 2	(0.001) 1	(0.041) 1	(0.001) 1	(0.017) 1	(0.003) 1
R42	9.155	**10.710**	4.966	**7.692**	**8.519**	**10.442**
	(0.057) 2	(0.005) 1	(0.083) 1	(0.021) 1	(0.014) 1	(0.005) 1
R54	**24.812**	**13.943**	5.957	**8.111**	**9.595**	**11.562**
	(0.000) 3	(0.001) 1	(0.051) 1	(0.017) 1	(0.008) 1	(0.003) 1
R66	**27.534**	**15.901**	8.949	**7.875**	**12.708**	**11.810**
	(0.000) 3	(0.000) 1	(0.062) 2	(0.019) 1	(0.002) 1	(0.003) 1
R78	**17.302**	**13.373**	**11.335**	**6.483**	**12.132**	**12.039**
	(0.002) 2	(0.001) 1	(0.023) 2	(0.039) 1	(0.002) 1	(0.002) 1
R90	**24.562**	**8.364**	**10.547**	4.463	**10.764**	**10.879**
	(0.000) 3	(0.015) 1	(0.032) 2	(0.107) 1	(0.005) 1	(0.004) 1
R102	**16.847**	**7.311**	3.439	4.690	**8.875**	**9.307**
	(0.010) 3	(0.026) 1	(0.179) 1	(0.096) 1	(0.012) 1	(0.010) 1
R114	**13.828**	**9.031**	4.140	**6.788**	**8.086**	**8.523**
	(0.032) 3	(0.011) 1	(0.126) 1	(0.034) 1	(0.018) 1	(0.014) 1

NOTE: Parentheses indicate significance levels; bold type indicates that the null hypothesis is rejected at the 5 percent significance level.

stationary. In all cases, following Bekaert and Hodrick (2001), the order of the VAR is determined by the Schwarz criterion.

Table 4 reports the LM statistic and the corresponding significance level for the tests on the VAR in levels. The lag length, chosen by the Schwarz criterion, is reported next to the significance level. Instances where the null hypothesis is rejected at the 5 percent level are in bold type. The results are reported using the 3-month *gensaki* rate and the 6-month Treasury rate as the short-term rate. For the entire sample period, the restrictions implied by the EH are frequently rejected at the 5 percent significance level and in every case at a slightly higher significance level. This finding may reflect evidence that Japanese rates are nonstationary, particularly for rates at the longer end of the maturity spectrum, since it appears that such rates are not cointegrated with the *gensaki* rate.

The EH does not fare well either for the period ending on October 1990 or for the period November 1990–June 1998, where the EH is frequently rejected at the 5 percent significance level and nearly always at the 10 percent level. Instances where the EH is not rejected at the 5 percent level when the evidence indicates that the rates are not cointegrated suggests the possibility that the test has low power when rates are not cointegrated.

The LM test is also applied to the VAR suggested by Campbell and Shiller (1987). Recall that this test is valid only if the rates are cointegrated with a cointegrating vector of (1, –1). Since interest rates are only cointegrated at the short end of the maturity spectrum and since the null hypothesis that the cointegrating vector is (1, –1) is frequently rejected, this test may be valid only for short maturities and only then for the first two sample periods.

The results for this test for the three sample periods are presented in Table 5. The results for the entire sample period are consistent with the results reported in Table 4. With one exception (the 42-month rate), the EH is rejected for the *gensaki* rate

Table 5

LM Statistics for Campbell-Shiller Test

	March 1981–January 2003		March 1981–October 1990		November 1990–June 1998	
	R3	R6	R3	R6	R3	R6
R6	**19.620** (0.001) 2	—	**7.396** (0.024) 1	—	**8.259** (0.016) 1	—
R18	**23.597** (0.001) 3	**6.762** (0.034) 1	4.671 (0.097) 1	4.565 (0.102) 1	0.300 (0.861) 1	2.540 (0.281) 1
R30	**12.193** (0.016) 2	**9.660** (0.008) 1	4.013 (0.134) 1	**13.551** (0.001) 1	0.712 (0.701) 1	1.783 (0.410) 1
R42	8.411 (0.078) 2	**6.851** (0.033) 1	3.144 (0.208) 1	**6.834** (0.033) 1	0.670 (0.715) 1	0.943 (0.624) 1
R54	**14.545** (0.006) 2	**7.636** (0.022) 1	**11.324** (0.023) 2	**6.086** (0.048) 1	0.644 (0.725) 1	1.807 (0.405) 1
R66	**22.301** (0.001) 3	**8.196** (0.017) 1	**9.947** (0.041) 2	**6.496** (0.039) 1	1.244 (0.537) 1	1.583 (0.453) 1
R78	**23.964** (0.001) 3	**7.481** (0.024) 1	**12.912** (0.012) 2	**6.107** (0.047) 1	0.693 (0.707) 1	1.430 (0.489) 1
R90	**14.680** (0.005) 2	4.272 (0.118) 1	3.188 (0.203) 1	3.536 (0.171) 1	0.316 (0.854) 1	0.663 (0.718) 1
R102	8.162 (0.086) 2	2.635 (0.268) 1	2.234 (0.327) 1	2.496 (0.287) 1	0.286 (0.867) 1	0.185 (0.911) 1
R114	7.627 (0.106) 2	5.165 (0.076) 1	2.941 (0.230) 1	4.664 (0.097) 1	0.437 (0.804) 1	0.693 (0.707) 1

NOTE: Parentheses indicate significance levels; bold type indicates that the null hypothesis is rejected at the 5 percent significance level.

at a very low significance level for long-term rates shorter than 102 months. The failure of the test to reject the EH when the long-term rate is 42 months and longer than 90 months is surprising and may be indicative of low power when rates are not cointegrated or the cointegrating vector is not (1, –1).

This interpretation is supported by the results for the March 1981–October 1990 and November 1990–June 1998 periods. When the *gensaki* rate is the short-term rate, the EH is rejected when the long-term rate is the 6-month rate. This is particularly true for the November 1990–June 1998 period, where the null hypothesis that the cointegrating vector is (1, –1) is always rejected regardless of the maturity of the long-term rate and the restrictions implied by the EH are never rejected. In any event, that the test tends to fail to reject the EH when rates appear not to be cointegrated or when the cointegrating vector appears to be different from (1, –1) suggests that this test may lack power when applied to data that do not satisfy the assumptions under which they are derived.

VII. IMPLICATIONS OF THE REJECTION OF THE EH FOR MONETARY POLICY

Finding that the EH does not hold presents a problem for the conventional view of the monetary policy transmission processes. According to this view, the central bank controls a very short-term interest rate and the effects of monetary policy are transmitted to longer-term rates in accordance with the EH. Since it is widely believed that investment spending depends on the behavior of relatively long-term interest rates, the fact that the EH appears not to hold for longer-term rates is problematic for the conventional view of monetary policy.

It is important to note, however, that the extent of this problem depends on exactly why the EH does not hold. One explanation for the failure of the EH—the overreaction hypothesis (ORH)—does not necessarily reduce the effectiveness of policy. Indeed, the efficacy of policy could be enhanced. According to the ORH, long-term rates overreact to expected changes in the short-term rate. Hence, during periods when the market expects interest rates to rise, long-term rates rise too much and too fast. Over time, and as expectations adjust, long-term rates fall while the short-term rate rises, which accounts for the failure of the EH. The ORH is not supported by evidence in the United States. Poole, Rasche, and Thornton (2002) show that the coefficient on a surprise change in the Federal Reserve Board's federal funds rate target for long-term rates is much smaller than that for short-term rates. Moreover, for rates longer than 12 months the estimated response is not statistically significant. Bekaert, Hodrick, and Marshall (2001) investigate a rational version of the ORH—namely, a "peso problem," where high interest rate regimes occur less frequently than rationally anticipated. They find that the peso problem cannot account for the failure of the EH in the United States.

Moreover, the ORH implies that long-term rates move more than short-term rates over the rate cycle. Thus, if the ORH is true, the variance of long-term rates should be generally larger than the variance of short-term rates. Figure 3 shows that this explanation is unlikely to account for the failure of the EH over the period January 1980–October 1990. The existence of the zero bound in the 1990s renders this explanation suspect since then.

Other explanations, such as the failure of rational expectations or, more generally, the market's inability to predict the behavior of interest rates, are more difficult for the conventional view of policy.[9] Either explanation implies that long-term rates need not be determined solely or in large part by the market's expectation for the policy rate. In such a circumstance, it is hard to understand how policy actions that affect very short-term rates are predictably transmitted along the yield curve. Some recent work using U.S. interest rates (Diebold and Li, 2003; Duffee, 2002; Carriero, Favero, and Kaminska, 2003; and Rudebusch, 2002) suggests that much of the failure of the EH in the United States might be due to the market participants' inability to forecast short-term rates.

One of the most frequently cited reasons for the failure of the EH is that the risk premium is time varying, rather than constant as the EH requires. One problem with this explanation is that any failure of the EH implies that the deviations in the risk premium are not *i.i.d.*, i.e., they are not time invariant. Hence, stating that the risk premium is time varying can be viewed as merely an alternative way of stating that the EH does not hold. For monetary policy to be effective, the actions of the monetary authority must be predictably transmitted to longer-term rates. For policy actions (which affect short-term rates) to have their desired effect on long-term rates, policymakers must be able to predict how the risk premium will vary over time—the efficacy of policy depends on policymakers' ability to predict changes in the risk premium.

While it is important to know that the EH does not hold, it is equally important to understand why it does not hold. It is now well established that, generally speaking, the EH does not hold in Japan. Research should now be focused on investigating why.

VIII. CONCLUSIONS

The deregulation of the Japanese bond market has generated interest in testing the expectations hypothesis (EH) of the term structure using Japanese data. This paper extends that literature by testing the EH for Japanese Treasury rates ranging in maturity from 6 to 114 months. This paper differs from previous tests of the EH using Japanese data in that it (i) considers the effects of nonstationarity on the EH, (ii) explicitly accounts for the stationarity of the data in testing the EH, and (iii) tests the EH by testing the restrictions imposed by it on two different VAR specifications of the short-term and long-term rates: one that assumes that interest rates are stationary and another that assumes that interest rates are nonstationary.

The results under the assumption that interest rates are stationary are not supportive of the EH. The EH is nearly always rejected at the 5 percent significance level, and in all but two instances rejected at the 10 percent significance level over the entire sample period and the subperiods considered.

The results are somewhat more supportive of the EH if one assumes that interest rates are nonstationary. A necessary condition for the EH holding is that the short-term and long-term rates are cointegrated. The evidence indicates that the *gensaki* rate is cointegrated with Treasury rates, but only for rates with maturities of 42 months or shorter. Consequently, the evidence suggests that the EH is likely to hold only at the short end of the maturity spectrum. Even in some of these instances, however, the hypothesis that the spread between the long-term and short-term rates is the equilibrium cointegrat-

[9] For example, Balduzzi, Bertola, and Foresi (1997) attempt to reconcile some of the empirical results by arguing that they are due in part to the market's failure to predict policy-induced short-term rate changes.

ing vector is rejected. In these instances, the EH is rejected.

The Campbell and Shiller (1987) test is applied to all combinations of short- and long-term rates, despite the fact that most rates do not satisfy the necessary conditions for the test to be applied. The EH is nearly always rejected at the 5 percent significance level over the entire sample period. This is not the case for the two shorter samples, however. Indeed, for about half of the combinations of short-term and long-term rates for the two shorter samples, the EH is not rejected, even at the 10 percent level. There are several issues that make this favorable interpretation for the EH problematic. First, and perhaps most troubling, the EH is rejected at a very low significance level when the short-term and long-term rates are the 3-month *gensaki* and 6-month Treasury rates, respectively. This is true for all three sample periods. Hence, the EH appears not to hold at the short end of the maturity spectrum, where most analysts (e.g., Rudebusch, 2002) believe that it is more likely to hold.

Second, the EH is frequently not rejected in cases where the evidence suggests (i) that interest rates are either not cointegrated or (ii) the spread between the long-term and short-term rates is not the cointegrating vector. Because the Campbell and Shiller (1987) test does not preserve the relationship between the levels of the rates if these conditions are not met, it is unclear whether the failure to reject the EH is because the EH holds or because the test has low power in such circumstances.

Third, relatively favorable results are obtained only when the *gensaki* rate is the short-term rate. When the 6-month rate is the short-term rate, the EH is frequently rejected at the shorter end of the term structure. Both this result and the rejection of the EH when the long-term rate is the 6-month rate could be due to idiosyncrasies in the behavior of the 6-month rate.

All in all, the EH appears not to fare well in Japan. If interest rates are nonstationary, the EH holds, at best, only at the short end of the maturity spectrum. This is encouraging, because most economists believe that the EH is likely to be the most relevant at the short end of the yield curve. This interpretation of the evidence presented here is consistent with recent analysis by Fujiki and Shiratsuka (2002) and Takeda and Yajima (2002), who find evidence that is broadly consistent with the EH using high-frequency, daily data over the period of the Bank of Japan's zero interest rate policy. Fujiki and Shiratsuka (2002) find that the yield curve flattens out over horizons of 3 months following the Bank of Japan's adoption of a zero interest rate policy and widens following the termination of the zero interest rate policy.[10]

REFERENCES

Balduzzi, Pierluigi; Bertola, Giuseppe and Foresi, Silverio. "A Model of Target Changes and the Term Structure of Interest Rates." *Journal of Monetary Economics*, July 1997, *39*(2), pp. 223-49.

Bekaert, Geert and Hodrick, Robert J. "Expectations Hypotheses Tests." *Journal of Finance*, August 2001, *56*(4), pp. 1357-94.

Bekaert, Geert; Hodrick, Robert J. and Marshall, David A. "'Peso Problem' Explanations for Term Structure Anomalies." *Journal of Monetary Economics*, October 2001, *48*(2), pp. 241-70.

Campbell, John Y. and Hamao, Yasushi. "The Interest Rate Process and the Term Structure of Interest Rates in Japan," in Kenneth J. Singleton, ed., *Japanese Monetary Policy*, NBER Monograph. Chicago: University of Chicago Press, 1993, pp. 95-120.

Campbell, John Y. and Shiller, Robert J. "Cointegration and Tests of Present Value Models." *Journal of Political Economy*, October 1987, *95*(5), pp. 1062-88.

Campbell, John Y. and Shiller, Robert J. "Yield Spreads and Interest Rate Movements: A Bird's Eye View." *Review of Economic Studies*, May 1991, *58*(3), pp. 495-514.

Carriero, Andrea; Favero, Carlo A. and Kaminska, Iryna. "Financial Factors, Macroeconomic Information and the Expectations Theory of the Term Structure of Interest Rates." Unpublished manuscript, 2003.

Chow, Gregory C. "A Comparison of Alternative Estimators for Simultaneous Equations." *Econometrica*, October 1964, *32*(4), pp. 532-53.

Dickey, David A. and Fuller, Wayne A. "Distribution of the Estimators for Autoregressive Time Series with a Unit Root." *Journal of the American Statistical Association*, June 1979, *74*(366), pp. 423-31.

[10] Takeda and Yajima (2002) perform tests using longer-term interest rates and generate much less convincing evidence in support of the EH.

Dickey, David A.; Jansen, Dennis W. and Thornton, Daniel L. "A Primer on Cointegration with an Application to Money and Income." Federal Reserve Bank of St. Louis *Review*, March/April 1991, *73*(2), pp. 58-78.

Diebold, Francis X. and Li, Canlin. "Forecasting the Term Structure of Government Bond Yields." NBER Working Paper No. 10048, National Bureau of Economic Research, October 2003.

Dittmar, Robert and Thornton, Daniel L. "Testing the Expectations Hypothesis: Some New Evidence." Working Paper No. 2003-021B, Federal Reserve Bank of St. Louis, 2003.

Duffee, Gregory R. "Term Premia and Interest Rate Forecasts in Affine Models." *Journal of Finance*, February 2002, *57*(1), pp. 405-43.

Engsted, Tom and Tanggaard, Carsten. "Cointegration in the U.S. Term Structure." *Journal of Banking and Finance*, January 1994, *18*(1), pp. 167-81.

Fujiki, Hiroshi and Shiratsuka, Shigenori. "Policy Duration Effect under the Zero Interest Rate Policy in 1999–2000: Evidence from Japan's Money Market Data." Institute for Monetary and Economic Studies, Bank of Japan *Monetary and Economic Studies*, January 2002, *20*(1), pp. 1-31.

Hall, Anthony D.; Anderson, Heather M. and Granger, Clive W.J. "A Cointegration Analysis of Treasury Bill Yields." *Review of Economics and Statistics*, February 1992, *74*(1), pp. 116-26.

Higo, Hideaki. "The Change in Liquidity in the Life Cycle of Japanese Government Securities," in *Market Liquidity: Research Findings and Selected Policy Implications*, CGFS Publication No. 11, Bank for International Settlements, 2000.

Hodder, James E. "The Cost of Capital for Industrial Firms in the U.S. and Japan," in William T. Ziemba, Waren Bailey, and Yasushi hamao, eds., *Japanese Financial Market Research (Contributions to Economic Analysis, No. 205)*. Amsterdam: North-Holland, 1991, pp. 571-93.

Kool, Clemens J.M. and Thornton, Daniel L. "A Note on the Expectations Hypothesis at the Founding of the Fed." Working Paper No. 2000-004C, Federal Reserve Bank of St. Louis, 2003 (forthcoming in *Journal of Banking and Finance*).

Leung, Kwok-Wai; Sanders, Anthony and Unal, Haluk. "The Structural Behavior of the Japanese Gensaki Rate," in *Japanese Financial Market Research (Contributions to Economic Analysis, No. 205)*. Amsterdam: North-Holland, 1991, pp. 557-68.

Okina, Kunio and Shiratsuka, Shigenori. "Policy Commitment and Expectation Formations: Japan's Experience under Zero Interest Rates." IMES Discussion Paper No. 2003-E-5, Institute for Monetary and Economic Studies, Bank of Japan, 2003.

Poole, William; Rasche, Robert H. and Thornton, Daniel L. "Market Anticipations of Monetary Policy Actions." Federal Reserve Bank of St. Louis *Review*, July/August 2002, *84*(4), pp. 65-93.

Rudebusch, Glenn D. "Term Structure Evidence on Interest Rate Smoothing and Monetary Policy Inertia." *Journal of Monetary Economics*, September 2002, *49*(6), pp. 1161-87.

Sarno, Lucio and Thornton, Daniel L. "The Dynamic Relationship between the Federal Funds Rate and the Treasury Bill Rate: An Empirical Investigation." *Journal of Banking and Finance*, June 2003, *27*(6), pp. 1079-110.

Shikano, Yoshiaki. "The Expectations Theory of the Term Structure of Interest Rates." Institute for Monetary and Economic Studies, Bank of Japan *Monetary and Economic Studies*, May 1985, *3*(1), pp. 47-70.

Shirakawa, Hiromichi. "Fluctuations in Yields on Bonds—A Reassessment of the Expectations Theory Based on Japanese and U.S. Data." Institute for Monetary and Economic Studies, Bank of Japan, *Monetary and Economic Studies*, September 1987, *5*(2), pp. 71-117.

Singleton, Kenneth J. "Interpreting Changes in the Volatility of Yields on Japanese Long-Term Bonds." Institute for Monetary and Economic Studies, Bank of Japan *Monetary and Economic Studies*, January 1990, *8*(1), pp. 49-77 (reprinted in *Japanese Financial Market Research*, 1991, pp. 527-56).

Stock, James H. and Watson, Mark W. "Testing for Common Trends." *Journal of the American Statistical Association*, December 1988, *83*(404), pp. 1097-107.

Takagi, Shinji. "Recent Developments in Japan's Bond and Money Markets." *Journal of the Japanese and International Economies*, 1988, *2*(1), pp. 63-91.

Takeda, Yosuke and Yajima, Yasuhide. "How the Japanese Government Bond Market Has Responded to the Zero Interest Rate Policy." *NLI Research*, February 2002.

Thornton, Daniel L. "Money Demand Dynamics: Some New Evidence." Federal Reserve Bank of St. Louis *Review*, March 1985, *67*(3), pp. 14-23.

Thornton, Daniel L. "The Conventional Test of the Expectations Hypothesis: Resolving the Anomalies When the Short-Term Rate Is the Federal Funds Rate." Working Paper 2000-003C, Federal Reserve Bank of St. Louis, 2002.

Thornton, Daniel L. "Tests of the Expectations Hypothesis: Resolving the Campbell-Shiller Paradox." Working Paper 2003-022b, Federal Reserve Bank of St. Louis, 2003a.

Thornton, Daniel L. "Monetary Policy Transparency: Transparent about What?" *Manchester School*, September, 2003b, *71*(5), pp. 478-97.

Asymmetric Effects of Monetary Policy in the United States

Morten O. Ravn and Martin Sola

This paper tests for the presence of asymmetric effects of monetary policy on aggregate activity using U.S. postwar quarterly data. We are interested in three types of asymmetry: (i) whether negative and positive monetary policy shocks have different effects on output; (ii) whether big or small shocks have different effects; and/or (iii) whether low-variance, negative shocks have asymmetric effects on output. We discuss the three possibilities below and explain under which conditions these asymmetries might take place.[1]

To date, the empirical literature has focused on a particular asymmetry that we call "the traditional Keynesian asymmetry," which states that positive monetary policy shocks have smaller real effects than negative monetary policy shocks—or, in a more extreme form, that only the latter shocks have real effects. This asymmetry can be derived under the assumption of either downward (upward) sticky (flexible) nominal wages or sticky prices together with rationing of demand.[2,3]

We also consider asymmetric effects that are implied by models with menu costs (see, among others, Ball and Romer, 1990, and Ball and Mankiw, 1994). In static (deterministic) settings, standard menu-cost models imply that "big" monetary policy shocks are neutral because firms would find it optimal to adjust nominal prices, while "small" monetary policy shocks would have real effects because keeping nominal prices fixed is associated with only a second-order cost. In other words, the firms have to decide—before the monetary policy shock is observed—whether to index their prices (at the cost of paying the menu cost) or not. Firms will choose indexation (which implies neutrality) only if the variance of monetary policy shocks is high. We extend the analysis by assuming that the monetary policy process can change between having a "high" variance and a "low" variance. This approach allows for identifying periods of neutrality and periods of non-neutrality.

Finally, we consider the case in which only small

1 Other types of asymmetric effects have been explored by Garcia and Schaller (1995), who examine whether monetary policy affects output differently in different phases of the business cycle, and Ravn and Sola (1997), who look at the effects of monetary policy on transitional dynamics. Hooker and Knetter (1996) analyze whether military procurement spending has asymmetric effects on employment, and they find that "big" negative shocks to procurement have proportionally larger effects on employment growth than large positive shocks or small shocks to procurement. Hooker (1996) examines whether there are asymmetries in the relationship between oil-price shocks and U.S. macroeconomic variables. He finds that the asymmetric effects in this relationship are fragile. Here we focus on the relationship between monetary policy shocks and aggregate activity. Lo and Piger (2003) examine regime switching in the response of U.S. output to monetary policy. They find evidence of such time variance and show that policy actions during recessions have larger output effects than policy actions during expansions.

2 Cover (1992) and DeLong and Summers (1988) have tested for this asymmetry in U.S. data: Cover (1992) finds firm support for the hypothesis in quarterly postwar data and shows that the results are robust to the specification of monetary policy and output. DeLong and Summers (1988) find that negative monetary policy shocks have a greater output effect than positive ones. Karras (1996) analyzes data for a number of European countries and finds strong evidence in favor of the traditional Keynesian asymmetry hypothesis. Parker and Rothman (2000) re-examine Cover's (1992) evidence for the pre-World War I and the interwar periods. They find that the type of asymmetry documented by Cover existed only during the latter stage of the Great Depression. Ravn and Sola (1996) show that, controlling for a regime change in monetary policy in 1979, the asymmetry documented by Cover is no longer significant.

3 Kandil (2002) explores the asymmetric effects of government spending and monetary shocks. Macklem, Paquet, and Phaneuf (1996) find results for Canada and the United States in line with those quoted above when including evidence from the yield curve and controlling for foreign factors. Sensier (1996) finds less-firm support for the asymmetry hypothesis in a study using U.K. data.

Morten O. Ravn is a professor at the London Business School and a research fellow at the Centre for Economic Policy Research. Martin Sola is a professor at the Universidad Torcuato di Tella and Birkbeck College, University of London. The authors thank Mark Hooker and Paul Evans for detailed comments and James Peery Cover for provision of data, as well as Charlie Bean, Michael Beeby, Paul Beaudry, John Driffill, Svend Hylleberg, Haris Psaradakis, Danny Quah, Ron Smith, Howard Wall, and seminar participants at the Bank of England, London School of Economics, University of Aarhus, Universitat Pompeu Fabra, Warwick University, the 1996 Econometric Society meeting in Istanbul, and the 1997 conference on Modelisation Macroeconometrique Dynamique at University Evry-Val d'Essonne, Paris.

Federal Reserve Bank of St. Louis *Review*, September/October 2004, *86*(5), pp. 41-60.

negative shocks to nominal demand affect real aggregate activity. Consider a dynamic menu-cost model in which there is positive steady-state inflation; firms can change prices costlessly every second period, but, if firms want to change prices in between the two periods, they must pay the menu cost. This gives rise to an asymmetric pricing rule in which "inaction" is optimal for a wider range of negative shocks than positive shocks. We call this case the "hybrid" asymmetry because it has similarities both to the traditional Keynesian asymmetry and to the menu-cost asymmetry.

To test for the asymmetric effects described above, we use a procedure that consists of estimating a monetary policy process that allows for changes in regime using the regime-switching model of Hamilton (1988) appropriately modified to our setting. We assume that the money supply is a regime-switching process that allows for changes in the mean and in the variance of the innovations to the process. This implies that we can distinguish between four different shocks to monetary policy: big positive shocks, big negative shocks, small positive shocks, and small negative shocks. The distinction between "big" and "small" here refers to the variance of the innovations in the two states.

This technique allows us to test for the existence of the three cases of asymmetric effects discussed above. We estimate a simultaneous system consisting of a monetary policy equation and an output equation, which includes the (change in the) current unanticipated shocks from the "monetary policy" relationship. We then test for asymmetries by introducing various parameter restrictions on the four different types of unanticipated monetary policy shocks in the output equation and by applying likelihood ratio tests.

We investigate two different sets of quarterly data for the U.S. postwar period. First we examine a data set for the period 1947-87, considered previously by Cover (1992), using M1 as the key monetary variable. The second set of data is for the period 1960-95, previously examined by, among others, Christiano, Eichenbaum, and Evans (1996). Here we measure monetary shocks on the basis of (the negative of) the federal funds rate.[4] The motivation for using the federal funds rate rather than M1 (or other money supply measures) is that the federal funds rate is widely recognized to be one of the primary monetary policy variables and is probably a more stable measure of monetary policy than M1.

Using the first set of data, we find that there have indeed been regime changes in the money-supply relationship. We find a low-growth, low-variance regime that spans the period from 1947 to around 1967 and a high-mean, high-variance regime that takes over for the majority of the period after 1968. When we test for the presence of asymmetric effects, we find that negative unanticipated money-supply shocks have greater real effects than positive unanticipated money-supply shocks.

Using the federal funds rate as the measure of monetary policy gives rise to different results. Again, the monetary policy process is divided into two regimes: one with a low mean and a low variance and another with a high mean and a high variance. The classification of the regimes is very different when M1 is used. The low-mean, low-variance regime occurs for most of the sample. The other regime dominates for a short period in the mid-1970s and the Volcker period. When we test for asymmetric effects using these alternative data, we find strong evidence in favor of the "hybrid asymmetry" (i.e., that only small negative monetary policy shocks have real effects). This finding is in line with the menu-cost model.[5]

The remainder of the paper is organized as follows. In the second section we look into the implications for asymmetric effects of standard menu-cost models. The third section is devoted to a description of the empirical method that we will apply. In the fourth section we examine the two alternative sets of U.S. data and test for the presence of asymmetries. In the fifth section we summarize and draw some conclusions.

THEORETICAL CONSIDERATIONS

First, to motivate and clarify the empirical analysis, we consider some of the theoretical possibilities for asymmetric effects of nominal demand on real output.

[4] We use the negative of the federal funds rate because a positive (negative) money-supply shock corresponds to a loosening (tightening) of monetary policy.

[5] Since we wrote this paper, a number of authors have examined these issues using slightly different techniques. Agénor (2001) examines the evidence on asymmetries using a vector autoregression (VAR) technique. He finds asymmetries for four emerging markets. Senda (2001) uses a panel technique to examine whether the degree of asymmetry is related to the magnitude of trend inflation and the variability of nominal gross domestic product (GDP) growth. Weise (1999) also applies a VAR technique focusing on asymmetries over the business cycle but also finds asymmetries in the response to money shocks of different sizes.

In the Keynesian literature building on sticky wages or sticky prices, the natural candidate for asymmetric effects is related to different real effects of positive and negative changes in nominal demand. Consider a model with downward (upward) sticky (flexible) nominal wages. Assume that the labor market initially clears at the nominal wage that corresponds to the price level (and expected price level) consistent with the current-level nominal demand and that the long-run supply curve is vertical. This implies that the supply curve will be vertical at the expected price level but positively sloped for price levels below the expected price level. Hence, unanticipated increases in nominal demand will be neutral, but unanticipated decreases in nominal demand will be associated with lower output and employment.

The problem with the analysis above is the lack of clear microeconomic foundations. Economic agents may adjust to the economic environment, and this can have implications for the result on asymmetric effects. Hence, it is important to consider models in which decision rules are explicitly derived. We will consider whether such asymmetries can arise in menu-cost-type models and derive the specific types of non-linearities in the relationship between activity and nominal demand.[6]

Here we follow the presentation in Ball and Romer (1990) and Ball and Mankiw (1994). Consider an economy with many price-setting agents, each of whom acts as a producer/consumer. Each agent produces a single differentiated good, which is sold at the nominal price, P_i. It is assumed that there is a small menu cost, denoted by $s > 0$, of changing nominal prices. Let the utility of agent i be given as

$$U_i = G\left(Y, \frac{P_i}{P}\right) - sD_i, \tag{1}$$

where Y denotes aggregate real spending, P is the aggregate price level, and D_i is a dummy variable that equals 1 if prices are changed and 0 otherwise. We assume that velocity is equal to unity, i.e., $Y = M/P$, where M denotes the nominal money stock. Equation (1) can then be written as

$$U_i = G\left(\frac{M}{P}, \frac{P_i}{P}\right) - sD_i. \tag{2}$$

In the absence of menu costs ($s = 0$), the first-order condition for each agent is that $G_2\left(\frac{M}{P}, \frac{P_i^*}{P}\right) = 0$, where G_2 denotes the derivative of G with respect to the second argument. In this case, in a symmetric equilibrium, changes in M are neutral. Such a symmetric equilibrium is assumed to exist and corresponds to $M = P = P_i = 1$.[7] Consider now an experiment where prices of all producers are set according to an expected money supply equal to 1, but after this $M \neq 1$ is realized. Each producer decides whether to pay the menu cost (setting prices equal to P_i^*), in which case money is neutral, or maintain prices (P_i), in which case money has real effects. Assume, first, that every price-setter except i expects all other price-setters not to change prices. The utility of not changing the price for agent i is then given as $U^{NA} = G(M,1)$. If the agent decides to change the price of good i, utility is given by $U^{CP} = G(M,P_i^*/P) - s$. Hence, inaction is an equilibrium if

$$U^{NA} - U^{CP} > 0 \Rightarrow G\left(M, \frac{P_i^*}{P}\right) - G(M,1) < s. \tag{3}$$

This condition implies that there is a range of money supplies for which inaction is a possible equilibrium.[8] Making a second-order Taylor approximation around $M = 1$, it can be shown that this range is given when

(4) M lies in the interval

$$\left(1 - M^*; 1 + M^*\right),\ M^* = \sqrt{\frac{-2G_{22}s}{G_{12}^2}}.$$

The range of money-supply shocks for which neutrality appears is given when

(5) M lies in the interval $(-\ ;M^{**})$ and/or $(M^{**};\ )$,

$$M^{**} = \sqrt{\frac{-2s}{G_{22}}}.$$

Thus, small money-supply changes have real effects when M lies in the interval $(1 - M^*; 1 + M^*)$; "big" changes are neutral when M lies in the interval $(-\ ;M^{**})$ and/or $(M^{**};\ )$. Hence, with menu costs and no other features, it is the *size* of the change in nominal demand that matters.

[6] Akerlof and Yellen (1985), Mankiw (1985), and Blanchard and Kiyotaki (1987) have analyzed how menu-cost models (or near-rationality) may affect the pricing decisions of firms and how this affects the real effects of changes in nominal demand.

[7] Strictly speaking, one also needs to assume that the second-order condition is fulfilled and that the equilibrium is stable (i.e., $G_{22}(1,1) < 0$ and $G_{12}(1,1) > 0$).

[8] It is possible that this range overlaps with a range of money supplies for which it is also optimal for all agents to change prices.

Above, the changes in money supply are zero-probability events. Alternatively, assume that money supply is a stochastic process with a mean M and a variance σ^2 and that agents must decide whether to pay the menu cost before observing the current money-supply shock. Thus, by construction, agents choose either indexation or non-indexation. Ball and Romer (1989, 1990) show that in this model non-indexation is an equilibrium for

$$(6)\qquad EG\left(\frac{M}{P_0},\frac{P_i^*}{P_0}\right)-EG\left(\frac{M}{P_0},1\right)\quad -\frac{G_{12}^2}{2G_{22}}\sigma^2<s,$$

where $1/P_0 \quad 1-\sigma^2G^2{}_{211}/(2G_{22})$. The difference between this case and the analysis above is that the decision of whether to pay the menu cost is determined by the *variance* of the money-supply shock. If the variance is high, money is neutral because firms perceive that there is a high probability of a big shock, while money has real effects if the variance is low. Thus, monetary policy is either *always* neutral or *always* non-neutral. This is a rather negative result since the theory as such does not have any testable (time-series) implications.

This latter implication can be overturned by a slight modification. Assume that the money supply can switch between two states of nature. In state i the variance of the money supply is σ_i^2 and $\sigma_1^2>\sigma_0^2$. Let us also assume that the state variable that dictates the variance of the money supply follows a first-order Markov process. Let π_{ij} be the probability that, given that the observed state today is i, the realized state tomorrow is j. The probability transition matrix is given by

$$(7)\qquad \Pi=\begin{bmatrix}\pi_{00} & \pi_{01}\\ \pi_{10} & \pi_{11}\end{bmatrix},$$

where each row sums to 1. Assume also that agents observe the current state when setting the initial price and when deciding whether to pay the menu cost or not. Then, using the same reasoning as above shows that inaction is an equilibrium when

$$(8)\qquad -\frac{G_{12}^2}{2G_{22}}\left(\pi_{00}\sigma_0^2+\pi_{01}\sigma_1^2\right)<s$$

when the current state is 0 and

$$-\frac{G_{12}^2}{2G_{22}}\left(\pi_{10}\sigma_0^2+\pi_{11}\sigma_1^2\right)<s$$

when the current state is 1.

There are two possible outcomes here. If (i) the difference between σ_0^2 and σ_1^2 is small or (ii) either π_{01} or π_{10} is close to 1, there will be either indexation or non-indexation in both states. If there is a non-trivial difference between the two variances and the states are relatively persistent, there will be indexation if today's state is 1 and non-indexation if today's state is 0. Hence, as in the standard menu-cost model, firms' actions depend on the monetary policy that they observe and their expectations of tomorrow's monetary policy.

Ball and Mankiw (1994) analyze a menu-cost model in which firms face a two-period problem and in which there is positive steady-state inflation (equal to $\dot{p}$). Each firm initially sets a price that can be changed next period, subject to a menu cost. They also assume the loss functions are quadratic such that, for a big enough menu cost, firms will choose a price that equals half the steady-state inflation rate in both periods.[9] If an unanticipated shock arrives in period 1, it might be optimal for firms to pay the menu cost and change prices. Since the optimal price in period 1 ($\dot{p}$) is already above the price set at period 0 ($\dot{p}/2$), it is clear that positive disturbances will lead to a greater incentive to change prices than negative disturbances. They show that in a quadratic setup, the range of non-action is given when M lies in the interval $\left(-\sqrt{s}-\dot{p}/2;\sqrt{s}-\dot{p}/2\right)$, which is symmetric around $-\dot{p}/2$ but asymmetric around 0. The model therefore implies an asymmetry that is similar to both the basic menu-cost results discussed above and to the traditional Keynesian asymmetry. We call this "hybrid" asymmetry.[10]

Finally, it is worth mentioning that imperfections in the labor market such as the existence of efficiency wage considerations or insider-outsider phenomena can be coupled with the menu-cost models. This has been investigated by Akerlof and Yellen (1985) and Ball and Romer (1990), and the

[9] If we let $\dot{p}$ denote the steady-state inflation rate, then with a quadratic loss function it is optimal to set prices at $\dot{p}/2$ in both periods, given that $s>\dot{p}^2/2$.

[10] Senda (2001) shows that the degree of asymmetry depends on the mean trend inflation rate and the variability of aggregate demand. Senda finds that the degree of asymmetry is non-trivially related to the mean inflation rate, increasing for low-to-moderate inflation rates but decreasing for high inflation rates. The reason for this is that, as inflation rates become very large, the cost of two-period price-setting becomes very large (in expected terms) and firms thus realize that they will probably want to change prices in the intermediate period. In this case, the asymmetry may become very small, although it still persists qualitatively. Senda also provides some favorable evidence of this hypothesis based on a panel of prewar and postwar data.

literature has shown that real rigidities increase the importance of nominal rigidities.

The cases discussed above relate to how different monetary policy shocks affect output. An alternative asymmetry is that monetary policy may affect aggregate activity differently during booms compared with recessions. Credit and liquidity may be readily available in booms, and it is likely that monetary shocks during these periods are neutral. In recessions, however, firms and consumers may find it harder to obtain funds and monetary policy might have real effects through the credit and liquidity channels. This is the mechanism examined in the research on financial market imperfections (see, e.g., Bernanke and Gertler, 1989, Gertler, 1992, Greenwald and Stiglitz, 1993, and Shleifer and Vishny, 1992). Although this possible asymmetry is of great interest, we shall not address it here but will concentrate on the above versions of asymmetric effects.

EMPIRICAL METHODOLOGY

In this section we describe our empirical methodology, which is related to the procedure used for testing the New Classical theories of information-based non-neutralities (developed by Lucas, 1972, 1975) in Barro (1977, 1978), Barro and Hercowitz (1980), Boschen and Grossman (1982), and, in particular, Mishkin (1982). Two relationships are estimated simultaneously. The first of these is a monetary policy relation from which one obtains estimates of the anticipated and unanticipated monetary policy shocks. These shocks then feed into an aggregate output equation. DeLong and Summers (1988) and Cover (1992) test whether positive and negative unanticipated monetary policy shocks have different effects on real activity and find strong support for the traditional Keynesian asymmetry in U.S. data.

Cover's (1992) methodology can be summarized as follows. First, one estimates simultaneously

$$\Delta m_t = \Phi(L)\Delta m_{t-1} + \Theta x_{t-1} + \varepsilon_t \tag{9}$$

and

$$\Delta y_t = \psi z_t + \beta^+\varepsilon_t^+ + \beta^-\varepsilon_t^- + \xi_t, \tag{10}$$

where Δ is the first-difference operator, m_t is the measure of the monetary policy, $\Phi(L)$ is a lag polynomial, Θ is a vector of parameters, x_{t-1} is a vector of predetermined regressors that reflects possible endogenous policy responses (and includes variables such as unemployment, changes in the monetary base, changes in output, government budget surpluses, changes in interest rates, and inflation), y_t is the measure of real aggregate activity, ψ is a parameter vector, z_t is a vector of regressors (which includes lagged changes in output and lagged changes in the Treasury bill rate), and ε_t^+ and ε_t^- are the positive and negative parts of ε_t from equation (9), defined as

$$\varepsilon_t^+ \equiv \max(0,\varepsilon_t),\ \varepsilon_t^- \equiv \min(0,\varepsilon_t). \tag{11}$$

Equation (9) is the monetary policy process and equation (10) is the aggregate output equation. The asymmetry hypothesis is a test of whether β^+ equals β^-; rejection of this restriction, together with β^+ being insignificantly different from zero and β^- significantly different from zero, supports the hypothesis.[11]

We extend this methodology along two lines. First, on the basis of the theory presented in the previous paragraph, we impose that monetary impulses have only temporary effects on the level of output. Because the output series we use here has a unit root, we stick to modeling the growth rate of output; but we change the specification of this equation[12] to

$$\Delta y_t = \psi z_t + \beta(e_t - e_{t-1}) + \xi_t, \tag{12}$$

where β is a vector of parameters and e_t is a vector of unanticipated money shocks specified later in the paper. This specification implies that any unanticipated shock associated with monetary policy will increase output only temporarily, exactly as stated in the theories that we have discussed.

Second, we differentiate not only positive and negative monetary policy shocks, but also big and small shocks. As made clear above, in a stochastic menu-cost model the relevant distinction between big and small is based on the variance of the unanticipated monetary policy shock. Hence, we estimate a monetary policy relationship that allows for this distinction, as a discrete-state regime-switching model.[13]

[11] Note that according to the specification of the money-supply equation and the output equation, money supply reacts to lagged variables, while output reacts to current monetary shocks. This assumption is contrary to standard assumptions made in the VAR literature but can be justified on the basis that the monetary authority may not have information on current output, while "true" real activity may be affected by actual current changes in monetary policy. We make this assumption mainly to make the analysis comparable to the previous contributions on asymmetries.

[12] We thank Paul D. Evans for pointing out the need to specify the system to account for the latter point.

[13] Such a technique has been used widely to characterize movements that arise when the moments of the variables under scrutiny change behavior over time; see, e.g., Hamilton (1988, 1989, 1990), Phillips (1991), Sola and Driffill (1994), and Ravn and Sola (1995). The basic elements of the method are described extensively in Hamilton (1994).

According to the regime-switching methodology, a time series is modeled as having discrete changes in its unconditional mean and/or variance and the changes in regime are dictated by an unobservable discrete-valued state variable, $s_t = 0,1$. We also add to the switching regression a set of conditioning variables that are not subject to regime changes. With this modification, we estimate a monetary policy equation that allows for changes in mean and variance. This leads us to the following specification:

$$(13)\quad \begin{aligned}&\left(\Delta m_t - \mu(s_t)\right)\\ &= \Phi(L)(\Delta m_{t-1} - \mu(s_{t-1})) + \Theta x'_{t-1} + \sigma(s_t)\eta_t ,\end{aligned}$$

where $\Phi(L)$ is a lag polynomial, Θ is a vector of parameters, x'_{t-1} is a vector of de-meaned predetermined variables (we include as regressors the log difference of non-borrowed reserves, the log difference of total reserves, the log difference of GDP, and the log difference of the implicit GDP deflator)[14] defined as $x - \mu_x$, $\mu(s_t)$ is a state-dependent mean, s_t is the discrete-valued state variable, and η_t is an i.i.d. $N(0,1)$ error term that is independent of s_t.

The monetary policy process can have two different means, μ_0 and μ_1, with associated variances σ_0^2 and σ_1^2. In the practical application these are estimated as $\mu_0 + \Delta\mu s_t$ and $\sigma_0 + \Delta\sigma s_t$. It is assumed that the (unobserved) states are generated by a two-state Markov process. Let π_{ij} be defined as $\pi_{ij} = P(s_t = i | s_{t-1} = j)$, $i,j = 0,1$. The probability transition matrix is given as

$$(14)\qquad \Pi = \begin{pmatrix} \pi_{00} & \pi_{01} \\ \pi_{10} & \pi_{11} \end{pmatrix},$$

where each of the transition probabilities is restricted to be non-negative and belongs to the unit interval.[15]

The division into big and small shocks is done as follows. Consider the expected money growth in period t, given information available at time t–1 and assuming momentarily that the information set includes the realization of the states. Expected money growth is given as

$$E^*_{t-1}\Delta m_t = \left(\mu_0 + \Phi\left(\Delta m_{t-1} - \left(\mu_0 + \Delta\mu\pi_{01}\right)\right) + \Theta x_{t-1}\right)$$

if $s_t = 0$ and

$$E^*_{t-1}\Delta m_t = \left(\mu_0 + \Delta\mu + \Phi\left(\Delta m_{t-1} - \left(\mu_0 + \Delta\mu\pi_{11}\right)\right) + \Theta x_{t-1}\right)$$

if $s_t = 1$, where * denotes that the information set includes the realized states. The unexpected monetary policy shocks in these two cases can then be defined as

$$\begin{aligned}\varepsilon_{0t} &= \Delta m_t \\ &-\begin{bmatrix}\mu_0 + \Phi \\ \left(\Delta m_{t-1} - \left(\mu_0 + \Delta\mu\pi_{01}\right)\right) + \Theta x_{t-1}\end{bmatrix} \sim N(0,\sigma_0^2)\\ \varepsilon_{1t} &= \Delta m_t \\ &-\begin{bmatrix}\mu_0 + \Delta\mu + \Phi \\ \left(\Delta m_{t-1} - \left(\mu_0 + \Delta\mu\pi_{11}\right)\right) + \Theta x_{t-1}\end{bmatrix} \sim N(0,\sigma_1^2)\ .\end{aligned}$$

The true information set, however, does not include the realized state, so we need to draw an inference on the regimes. To do this we use the estimates of the probabilities of being in each of the two regimes. Let $P(s_t = i | I_t)$ be the (estimated) probability conditional on information available at time t that the state is equal to i at time t using the (modified) Hamilton filter. Assume also that state 0 is the state in which the variance of unanticipated monetary policy shocks is low. We can then define the two shocks in the following manner:

$$(15)\quad \begin{aligned}e_t^S \equiv &\left(\Delta m_t - \left[\mu_0 + \Phi\begin{pmatrix}\Delta m_{t-1} - \\ \left(\mu_0 + \Delta\mu\pi_{01}\right)\end{pmatrix} + \Theta x_{t-1}\right]\right)\\ &\times P(s_t = 0 | I_t)\end{aligned}$$

and

$$(16)\quad \begin{aligned}e_t^B \equiv &\left(\Delta m_t - \left[\mu_0 + \Delta\mu + \Phi\begin{pmatrix}\Delta m_{t-1} - \\ \left(\mu_0 + \Delta\mu\pi_{11}\right)\end{pmatrix} + \Theta x_{t-1}\right]\right)\\ &\times P(s_t = 1 | I_t)\ .\end{aligned}$$

Next, each of these two shocks can be divided into their positive and negative parts, which we denote by + (positive) and – (negative), using the same technique as in the previous section. Accordingly, we end up with four monetary policy shocks, $e_t = \{e_t^{B+}, e_t^{B-}, e_t^{S+}, e_t^{S-}\}$. This construction allows us

---

[14] We de-mean the non-switching exogenous variables so that $\mu(s_t)$ can be interpreted as the unconditional mean of money growth.

[15] Note that we do not allow for regime switching in the exogenous variables. To allow these variables to have changes in regime will require imposing either that they all switch simultaneously with the money supply (see, e.g., Sola and Driffill, 1994) or that each variable is allowed to switch independently (see, e.g., Ravn and Sola, 1995). The first approach is applicable when the variables are closely related (for example, for interest rates of bonds of different maturities), but does not naturally occur in the present analysis. The second approach has the disadvantage that the increase in the number of states quickly makes it intractable.

to test for the presence of asymmetric effects using the following procedure.[16]

We estimate jointly the monetary policy equation (13) and the following version of the output equation[17]:

$$
\begin{aligned}
(17)\qquad \Delta y_t &= \psi z_t + \beta^{B+}\Delta e_t^{B+} + \beta^{B-}\Delta e_t^{B-} \\
&\quad + \beta^{S+}\Delta e_t^{S+} + \beta^{S-}\Delta e_t^{S-} + \xi_t .
\end{aligned}
$$

First we estimate equations (13) and (17), imposing that all the β coefficients are equal to 0—that is, that money has no real effects. We call this Case 0. Next, we estimate the system that allows the unanticipated monetary policy shocks to enter unrestricted. We call this Case 1. At this point one can look at the significance of each of the shocks as a check on signs of asymmetric effects; one can also check for monetary neutrality by using a likelihood-ratio (LR) test (with four degrees of freedom) if it is tested against Case 0.

The tests for asymmetries are carried out in a sequential manner using LR tests by imposing parameter restrictions on the coefficients on Δe_t. First we impose the following:

(18) **Case 2:** $H_0{:}\beta^{B+} = \beta^{B-} = \beta^{S+} = \beta^{S-}$.

Asking whether Case 2 is a valid simplification of Case 1 is equivalent to testing for the absence of any asymmetry and can be performed as an LR test that is χ^2–distributed with three degrees of freedom under the null. If these restrictions are rejected, the tests for the two versions of asymmetric effects are carried out by imposing a number of different parameter restrictions.

First, consider the case of testing for the asymmetry hypothesis that positive and negative monetary policy shocks have different effects; this can be tested in two steps. According to this hypothesis, it should not matter whether a given monetary policy shock is big or small. Hence, we impose the following:

(19) **Case 3:** $H_0{:}\beta^{B+} = \beta^{S+}$ and $\beta^{B-} = \beta^{S-}$.

Comparing Case 3 with Case 1 constitutes the first assessment of this hypothesis. It is further required that positive shocks are neutral. Hence, we impose the following:

(20) **Case 4:** $H_0{:}\beta^{B+} = \beta^{S+} = 0$ and $\beta^{B-} = \beta^{S-}$.

Comparing Case 4 with Case 3 is a way to assess whether positive shocks are neutral. If these tests are passed and the coefficient on the negative shocks is significantly positive, the data support the traditional Keynesian asymmetry hypothesis.

The other asymmetry hypothesis can be tested similarly. First we impose the following:

(21) **Case 5:** $H_0{:}\beta^{B+} = \beta^{B-}$ and $\beta^{S+} = \beta^{S-}$.

Testing Case 5 against Case 1 constitutes the first part of the hypothesis. The second part imposes that big shocks are neutral:

(22) **Case 6:** $H_0{:}\beta^{B+} = \beta^{B-} = 0$ and $\beta^{S+} = \beta^{S-}$.

Again, we test this specification against Case 5, and if the test is passed the hypothesis is backed by the data. A last case to consider is the hybrid version in which only small negative shocks have real effects. We can test the hybrid version in any of the sequences outlined above. This can be performed by imposing the following:

(23) **Case 7:** $H_0{:}\beta^{B+} = \beta^{B-} = \beta^{S+} = 0$.

We test this against Case 1 because this case might not be nested within Case 4 or Case 6 if the null is correct.[18]

EMPIRICAL TESTS FOR THE UNITED STATES

In this section, we empirically test for the different varieties of asymmetric effects of nominal demand on real activity discussed in the previous section. We look at two alternative sets of quarterly data for the United States.[19] The first data set covers the period 1948-87 and the second data set covers the period 1960-95.

In both applications we use the empirical method described above, but the two applications differ in the measure of monetary policy that is used.

[16] Alternatively, one can use a method of simulated moments to obtain estimates of the unexpected money growth.

[17] The models were estimated using a maximum-likelihood estimator for the joint system of equations.

[18] The main difference between our approach and other applications is the definition of big and small shocks. Demery (1993) makes the distinction by defining the former as those that are in a two-standard-error interval around 0 and the latter as those not belonging to this interval (i.e., as outliers). Caballero and Engel (1993) apply a similar strategy when testing for asymmetries in the price adjustments of firms in the face of nominal rigidities and changes in demand. This definition is not appropriate in light of our analysis in the previous section and may produce estimates of wrongly identified monetary policy shocks.

[19] The data are described in more detail in the appendix.

For the first data set, we use the logarithm of M1 as the measure of monetary policy. When looking at the more recent data, we use the (negative of the) federal funds rate. The procedure we use depends on the time-series properties of the data in question, since we need to forecast the monetary policy process. Preliminary data analysis revealed that the logarithm of the money supply has a unit root, whereas the federal funds rate is a stationary process. Given this, for the federal funds rate we estimate equation (13) in levels using $-r^f$ as the measure of Δm_t. For the money-stock measure, we use the first difference of the logarithm to construct the one-step-ahead forecast error. Let M_t be the money series that has a unit root—then the one-step-ahead forecast will be

(24) $$M^e_{t+1} = M_t + \ (L)\Delta M_t ,$$

where $\hat{\ }(L)$ is an estimator of $\ (L)$,

(25) $$\widehat{\Delta M}_t = \ (L)\Delta M_{t-1} .$$

This implies that the unanticipated money shock can be written as

(26) $$e_{t+1} = M_{t+1} - \left(M_t + \ (L) \ M_t\right).$$

Note that the theoretical relationships that we want to consider are in "levels," whereas the empirical model is in "differences"; therefore, we consider relationships of the type expressed in equation (12) to preserve the theoretical structure of interest:

(27) $$\Delta y_t = \psi z_t + \beta\left(e_t - e_{t-1}\right) + \xi_t .$$

Another issue to be addressed is the specification of the relationships for monetary policy and output. The money-supply process is specified as in Barro and Rush (1980), as in Mishkin (1982), or as an "optimal" money supply.[20] The two specifications of the output equation have in common that z_t includes a constant, lagged change in real output, ε_t^+ and ε_t^-; but the two specifications differ in whether the change in the T-bill rate is included or not.

In all of the above money-supply processes there are signs (i) of misspecification related to the existence of an outlier (at the first quarter of 1983, when the Volcker regime ended) in the money-supply residuals and (ii) of heteroskedasticity of the money-supply residuals (details can be found in Ravn and Sola, 1996). For these reasons, we estimate (using a general-to-specific approach) an alternative money-supply process that includes the first lag of M1 growth; the fourth- to the sixth-quarter lags of the (log of the) federal government's budget surplus; the first, fifth, and sixth lags of the log difference of the monetary base; the two-quarter lag of the unemployment rate; the second and the sixth lags of output growth; and the first, third, and fifth lags of the first difference of the T-bill rate.[21,22] This relationship was identified by testing downward from a relationship that initially included six lags of all the variables.

In the second application we use the negative of the federal funds rate as the measure of monetary policy. We use the negative of the federal funds rate such that a positive shock to the monetary policy process can be interpreted as a loosening of monetary policy. In this application, the vector of regressors in the monetary policy relationship includes four lags of the (negative of the) federal funds rate, four lags of the log difference of GDP, four lags of the log difference of nonborrowed reserves, four lags of the log difference of total reserves, and four lags of the log difference of the implicit GDP deflator.[23]

For both sets of data we specified the output equation (17) such that it includes one lag of output growth, the first difference of the T-bill rate, and the lag of the first-differenced T-bill rate.

Results for M1

Single-Equation Estimates of the Money Supply. We first turn to the results of single-equation estimates of the money-supply process with changes in regime. Figure 1 illustrates the first difference of the log of M1, and Table 1 reports

[20] In the Barro-Rush specification, the vector of regressors x_{t-1} includes a constant, the unemployment rate, and the contemporaneous real federal expenditures to normal expenditures. In the "modified Mishkin" specification, x_{t-1} contains constant, lagged changes in money supply, lagged changes in the T-bill rate, and lagged values of the federal government's budget surplus. The "optimal" specification includes various elements of the above variables as well as lagged values of the changes in the monetary base.

[21] It also turns out that, once one corrects for the presence of the outlier, the results on asymmetric effects are no longer valid. Specifically, one can in this case no longer reject a hypothesis that the positive and negative shocks have the same effect on output and that they are neutral. Details on results are given in Ravn and Sola (1996, Table 3).

[22] Belongia (1996) documents another problem with the asymmetry result for the U.S. data. Belongia (1996) shows that if one uses a divisia index for the money stock, then one cannot reject the hypothesis that positive and negative money-supply shocks have symmetric effects.

[23] We also experimented with including the unemployment rate, but this variable did not affect the results. It should also be noted that we have not included the index of sensitive commodity prices that Christiano, Eichenbaum, and Evans (1996) introduce to address the "price puzzle." This issue is not important for our analysis.

Figure 1

The Growth Rate of M1

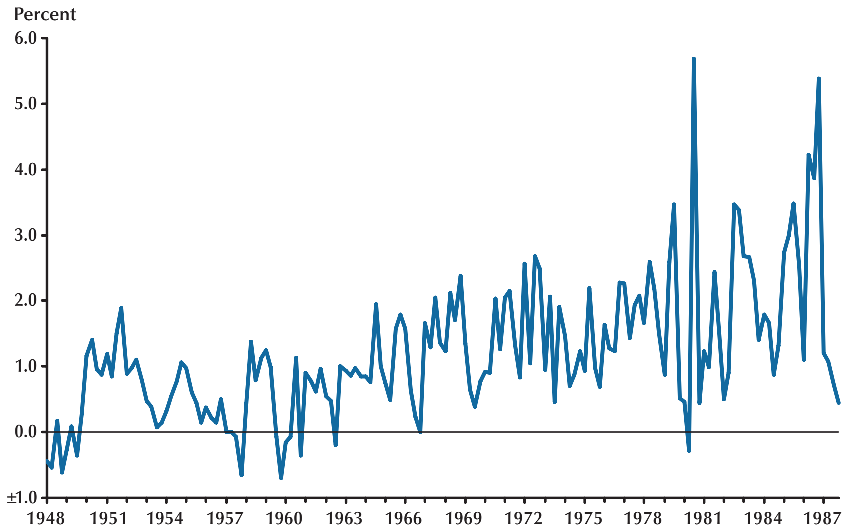
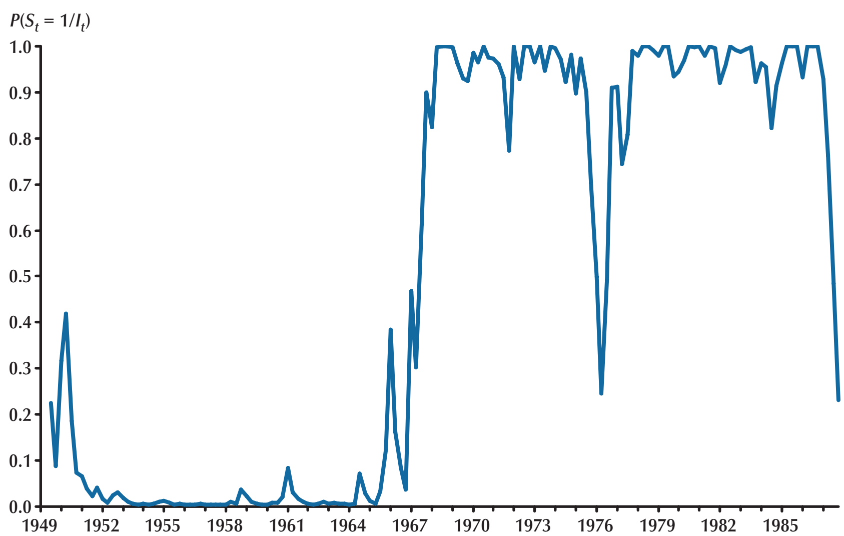


Figure 2

Filter Probabilities: Single Equation Results for M1



the results for the estimation of the money-supply process with changes in regime. We find that the changes in both the mean and the variance of the process are significant. The estimates suggest that there is a low-mean, low-variance regime where the mean is around 0.7 percent per quarter and the standard deviation around 0.4 percent and a high-mean, high-variance regime where the mean is around 1.65 percent per quarter and the standard deviation around 0.8. That is, the mean and standard deviation of the innovation in the "high" state (state 1) of money supply are roughly twice the corresponding numbers in the "low" state (state 0). Note also that both regimes are quite persistent, since the diagonal elements of the transition matrix are both in excess of 0.98.

Table 1

Money Supply with Changes in Regime: M1 Process, Single-Equation Results, 1948:Q1–1987:Q4

Variable	Estimate	Variable	Estimate	Test statistic
$\Delta m1_{t-1}$	0.257 (0.072)	μ_0	0.726 (0.131)	$Q(1) = 0.295$ [0.587]
u_{t-2}	0.403 (0.189)	$\Delta\mu$	0.934 (0.235)	$Q(10) = 9.521$ [0.709]
fs_{t-4}	–1.160 (0.608)	σ_0	0.417 (0.041)	$QQ(1) = 2.264$ [0.132]
fs_{t-5}	2.254 (0.843)	$\Delta\sigma$	0.368 (0.077)	$QQ(10) = 16.120$ [0.096]
fs_{t-6}	–1.234 (0.528)	π_{00}	0.985 (0.015)	
Δb_{t-1}	0.078 (0.092)	π_{11}	0.989 (0.013)	
Δb_{t-5}	0.141 (0.079)			
Δb_{t-6}	–0.180 (0.089)			
Δy_{t-2}	0.144 (0.042)			
Δy_{t-6}	0.115 (0.048)			
Δtbr_{t-1}	–0.405 (0.054)			
Δtbr_{t-3}	–0.199 (0.058)			
Δtbr_{t-5}	–0.144 (0.056)			

NOTE: $\Delta m1$ is the log difference of M1; u is the unemployment rate; fs is the log of the federal government's budget surplus; Δb is the log difference of the monetary base; Δy is the log difference of GNP; and Δtbr is the difference of the T-bill rate. $Q(x)$ ($QQ(x)$) is the Box-Pierce test for autocorrelation in the standardized residuals (squared standardized residuals) of order x. Numbers in parentheses are standard errors; numbers in brackets are probabilities.

Figure 2 illustrates the estimated probabilities of being in regime 1, the regime in which the mean and the variance are both high. The filter divides the sample very clearly into the two regimes, and the estimates imply that money growth and the variance of money growth were low from the start of the sample until 1967. From 1967 to 1987:Q4, the probability of being in the regime with high mean and high variance is practically equal to 1, with the exception of the last three quarters of 1976 and the final three observations. It should also be noted that there are no signs of specification errors in the regression residuals.

Tests for Asymmetric Effects. We first estimate Case 0, that is, the output equation, without money entering into it. This is reported in the first column of Table 2. We see that the output equation is relatively well estimated, with no signs of misspecification in the errors. Next, we estimate the system, letting each of the money-shock components enter unrestricted (Case 1). None of the four money-supply shocks are significant individually, but the LR test implies that the four shocks are significant jointly. There is no clear pattern that leads one to suspect the presence of asymmetries, but to test this more formally we use the procedure outlined above.

Table 2

Output Equation ML Estimates: M1 Measure, 1948:Q1–1987:Q4

Variable	Case 0	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7
Constant	0.540	0.502	0.520	0.502	0.520	0.507	0.502	0.533
	(0.098)	(0.098)	(0.097)	(0.098)	(0.097)	(0.098)	(0.096)	(0.098)
Δy_{t-1}	0.329	0.370	0.353	0.369	0.355	0.371	0.376	0.337
	(0.075)	(0.076)	(0.074)	(0.075)	(0.075)	(0.055)	(0.075)	(0.075)
Δtbr_t	0.272	0.278	0.284	0.278	0.277	0.285	0.285	0.272
	(0.067)	(0.067)	(0.067)	(0.067)	(0.067)	(0.067)	(0.067)	(0.068)
Δtbr_{t-1}	0.081	0.051	0.071	0.052	0.057	0.070	0.071	0.079
	(0.071)	(0.073)	(0.071)	(0.073)	(0.071)	(0.071)	(0.070)	(0.071)
Δe_t^{S+}	—	0.071	0.184	0.025	—	0.395	0.448	—
		(0.391)	(0.083)	(0.160)		(0.219)	(0.174)	
Δe_t^{S-}	—	0.370	0.184	0.415	0.363	0.395	0.448	0.335
		(0.386)	(0.083)	(0.220)	(0.166)	(0.219)	(0.174)	(0.377)
Δe_t^{B+}	—	0.003	0.184	0.025	—	0.075	—	—
		(0.106)	(0.083)	(0.160)		(0.178)		
Δe_t^{B-}	—	0.432	0.184	0.415	0.363	0.075	—	—
		(0.240)	(0.083)	(0.220)	(0.166)	(0.178)		
σ_Y	0.972	0.950	0.956	0.950	0.954	0.950	0.947	0.969
$Q(1)$	0.174	0.371	0.169	0.392	0.232	0.186	0.191	0.269
	[0.677]	[0.543]	[0.680]	[0.531]	[0.630]	[0.666]	[0.662]	[0.604]
$Q(10)$	8.474	7.652	7.957	8.050	8.408	7.487	8.997	9.518
	[0.583]	[0.663]	[0.633]	[0.624]	[0.589]	[0.679]	[0.532]	[0.484]
$QQ(1)$	2.165	4.964	3.166	4.991	4.312	3.441	2.982	2.932
	[0.141]	[0.026]	[0.075]	[0.026]	[0.038]	[0.064]	[0.084]	[0.087]
$QQ(10)$	7.306	12.575	9.958	12.427	11.246	11.114	10.867	8.264
	[0.696]	[0.248]	[0.444]	[0.258]	[0.339]	[0.349]	[0.368]	[0.603]
Log likelihood	–358.71	–352.78	–356.32	–352.79	–356.42	–356.00	–356.05	–358.33
LR test		11.94[0)]	7.08[1)]	0.03[2)]	7.25[3)]	6.44[4)]	0.098[5)]	11.10[6)]
		[0.018]	[0.069]	[0.983]	[0.007]	[0.040]	[75.42]	[0.011]

NOTE: See note to Table 1. The monetary shocks refer to unanticipated shocks. We do not report the estimates of the money-supply equations (which are jointly estimated by ML), but they are available upon request. 0) *LR* test of Case 0 vs. Case 1; 1) *LR* test of Case 2 vs. Case 1; 2) *LR* test of Case 3 vs. Case 1; 3) *LR* test of Case 4 vs. Case 3; 4) *LR* test of Case 5 vs. Case 1; 5) *LR* test of Case 6 vs. Case 5; 6) *LR* test of Case 7 vs. Case 6.

First we impose the restrictions based on Case 2, that is, absence of asymmetries. We obtain a *p*-value of 6.94 percent for this hypothesis, implying that there is no strong evidence in favor of asymmetric effects once one allows all the coefficients to enter unrestricted in the alternative hypothesis. Notice also that, once these restrictions are imposed, we obtain significant coefficients on the unanticipated shocks to M_1. However, we still test for asymmetries and first look at the traditional Keynesian hypothesis; that is, we impose the restrictions of Case 3. These restrictions imply that big and small shocks enter with the same coefficients. The parameter estimates now imply that negative shocks enter with a coefficient that is much larger than that of positive shocks. Furthermore, the LR test indicates that the restrictions cannot be rejected at any conventional significance level (in spite of the coefficients being insignificant individually).

When we impose that the positive shocks are neutral, we find that the negative shocks become significant; but, when tested against Case 3, the restrictions of Case 4 are strongly rejected. It should be noted, however, that the positive shocks have very small effects on output. Thus, even though we formally reject that these shocks are neutral, their quantitative effects appear limited.

The other alternative to be tested is whether big and small money-supply shocks have asymmetric effects on output. First we impose the restrictions under Case 5 (i.e., that it is irrelevant whether the shocks are positive or negative). We find that the probability value of the LR test of this hypothesis is 4 percent, which implies that we would reject the null hypothesis at the 5 percent level. One might be tempted to continue with the hypothesis, given the marginal rejection. In that case one would not be able to reject that big shocks are neutral, thus finding evidence in favor of the menu-cost type of asymmetry. However, the likelihood of Case 5 is much worse than the competing likelihood of Case 3: Thus, in this respect, Case 3 appears to be the better specification. Finally, we need to look at Case 7, the case based on the hybrid asymmetry. This case is rejected regardless of which alternative it is tested against.

In conclusion, the data give some support to the idea that negative monetary policy shocks have larger real effects than positive monetary policy shocks. However, at the same time, we cannot formally reject that all types of shocks have identical effects on output—that is, that monetary policy has symmetric effects. And, regardless of this, we find very small monetary policy effects. Thus, while the evidence does not directly contradict previous evidence, the results do not strongly support the traditional Keynesian asymmetry.

Results for the Federal Funds Rate

As discussed previously, there are reasons to expect that the results above might be hampered by the structural instability of M1 demand. It has previously been shown that M1 demand has been relatively unstable in the 1980s and the 1990s. This implies that the shocks identified above, as a "monetary policy" shock, may well indeed be a mixture of money-demand and money-supply shocks. (See, e.g., Baba, Hendry, and Starr, 1992, or Stock and Watson, 1993, for a discussion.) It has also been claimed that the federal funds rate may be a better indicator of monetary policy.[24] The reason is that much of the Federal Reserve's intervention takes place in the form of changes in nonborrowed reserves, which affect the interest rate in the reserve market, that is, the federal funds rate. For these reasons we now take up the question of asymmetric effects using the federal funds rate rather than M1. To facilitate an easy comparison with the analysis for M1, we will transform the federal funds rate and measure it by the negative of the federal funds rate such that positive shocks indicate a loosening of monetary policy. The federal funds rate is illustrated graphically in Figure 3. One notices immediately the volatile behavior of the federal funds rate in the early 1980s.

Single-Equation Estimates. We start by looking at the results of single-equation estimates of the federal funds rate process using the regime-switching technique. The federal funds rate process includes four lags of the following five variables: (i) the federal funds rate, (ii) the log-difference of GDP, (iii) the log-difference of the implicit GDP deflator, (iv) the log-difference of non-borrowed reserves, and (v) the log-difference of total reserves.[25]

Table 3 reports the single-equation results of the estimation of the process for the federal funds rate. As for M1, we find that there are clear signs of changes in regime. We find a low-mean, low-variance regime and a high-mean, high-variance regime. In the low regime the mean of the federal funds rate is estimated to be around 6.4 percent and the standard deviation to be 0.42 percent. In the high regime, the mean is estimated to be around 8.4 percent and the standard deviation to be 2.2 percent. Evidently, it is the change in the variance that dominates the change in regime in this process. Furthermore, from the estimates of the Markov transition probabilities, one can see that the low-mean, low-variance regime is much more persistent than the high-mean, high-variance regime. The probabilities imply that the expected duration of the low-mean, low-variance regime is close to 15 years, while the expected duration of the high-mean, high-variance regime is exactly equal to 2 years.

Figure 4 illustrates the estimated probabilities of each of the two regimes. The regime with low funds rates and a low variance of the innovations is estimated to dominate most of the sample period. There are two periods in which the regime with high funds rates and high volatility takes over. The first period is the period immediately after the first oil-

[24] Hamilton (1996) provides an excellent discussion and analysis of the federal funds daily market.

[25] The results are robust to changes in the federal funds rate process. We experimented with the inclusion of the unemployment rate, with using the CPI rather than the GDP deflator, and with using industrial production rather than GDP. We also experimented with alternative lag lengths and got the same results as those reported here.

Figure 3

Federal Funds Rate

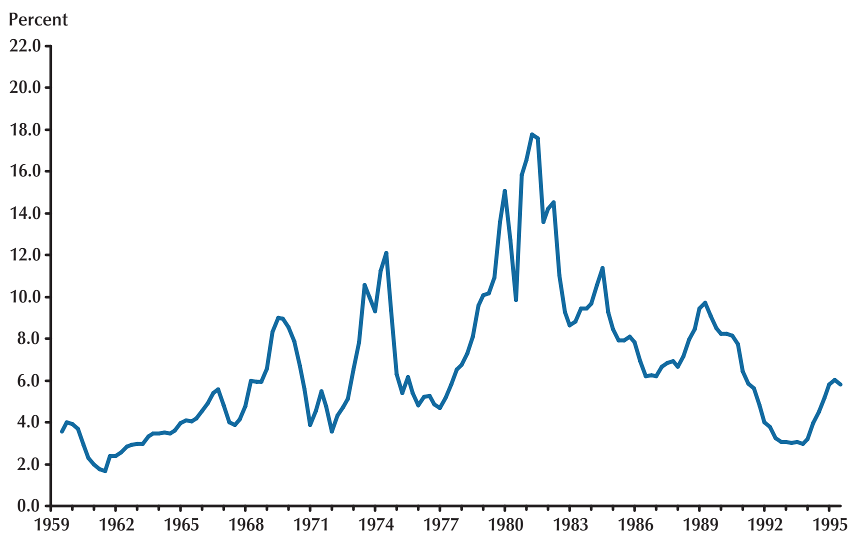
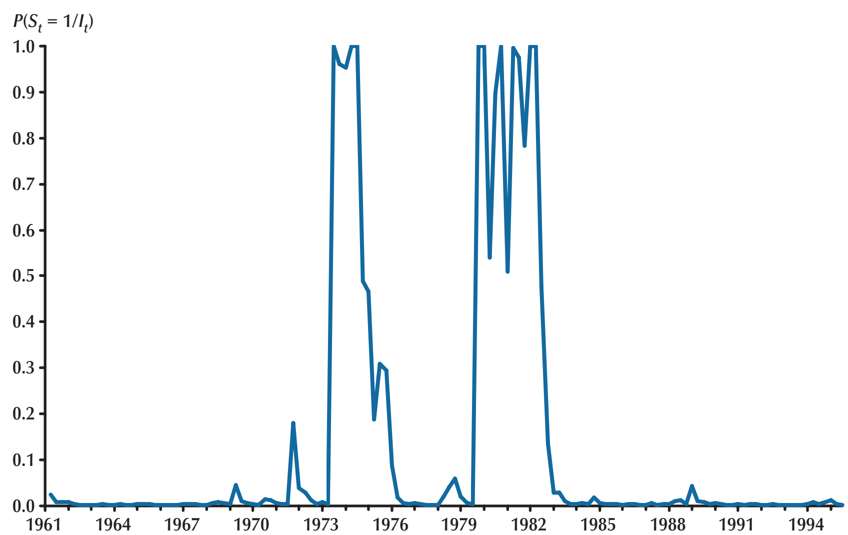


Figure 4

Filter Probabilities: Single-Equation Results for the Funds Rate



price shock, 1973:Q3–1975:Q4. The second period is, not surprisingly, the Volcker period, 1979:Q3–1982:Q3. (One might also include 1982:Q4 in this regime, but our estimates imply that the probability of the high regime is 13.4 percent for this observation.) These results seem much more sensible than the dating of regimes in the application using M1.

Tests for Asymmetric Effects Using the Federal Funds Rate. In this application we use GDP as the measure rather than gross national product, which is used for the analysis with the M1 data. The results are reported in Table 4.

In the first column we report the results for the output equation that excludes the monetary policy shocks; in the second column, we report the results when each of the four shocks enter unrestricted.

Table 3

Money Supply with Changes in Regime: Federal Funds Rate Process Single-Equation Results, 1959:Q3–1995:Q3

Variable	Estimate	Variable	Estimate	Test statistic
$-ff_{t-1}$	1.226	Δtr_{t-1}	–0.178	$Q(1) = 0.002$
	(0.108)		(0.044)	[0.968]
$-ff_{t-2}$	–0.283	Δtr_{t-2}	0.147	$Q(10) = 7.612$
	(0.180)		(0.053)	[0.667]
$-ff_{t-3}$	0.049	Δtr_{t-3}	–0.011	$QQ(1) = 0.305$
	(0.163)		(0.059)	[0.581]
$-ff_{t-4}$	–0.052		–0.024	$QQ(10) = 7.378$
	(0.093)	Δtr_{t-4}	(0.046)	[0.689]
Δnbr_{t-1}	0.145			
	(0.038)			
Δnbr_{t-2}	–0.113	μ_0	–6.362	
	(0.043)		(0.675)	
Δnbr_{t-3}	–0.041	$\Delta\mu$	–2.031	
	(0.045)		(0.486)	
Δnbr_{t-4}	0.047	σ_0	0.418	
	(0.037)		(0.027)	
Δy_{t-1}	–0.173	$\Delta\sigma$	1.760	
	(0.057)		(0.390)	
Δy_{t-2}	–0.053	π_{00}	0.875	
	(0.055)		(0.080)	
Δy_{t-3}	–0.069	π_{11}	0.984	
	(0.054)		(0.011)	
Δy_{t-4}	0.021			
	(0.051)			
Δp_{t-1}	–0.377			
	(0.172)			
Δp_{t-2}	–0.285			
	(0.162)			
Δp_{t-3}	0.217			
	(0.168)			
Δp_{t-4}	0.207			
	(0.168)			

NOTE: $-ff$ is the negative of the federal funds rate; Δnbr is the log difference of non-borrowed reserves; Δy is the log difference of GDP; Δp is the log difference of the implicit GDP deflator; and Δtr is the log difference of total reserves.

For this specification, only the big negative shock enters significantly and the other components enter with negative coefficients, although they are insignificantly different from 0. When tested against Case 0, we strongly reject that money is neutral. From this perspective, there are signs of asymmetries, but the negative point estimates on some of the shocks seem slightly puzzling. In column 3 we impose that all four shocks enter with identical coefficients, and, again, the LR test strongly rejects this specification. Thus, we proceed to test for either of the two asymmetry hypotheses.

First we impose the parameter restrictions for Case 3. These restrictions have a probability value of just above 1 percent and are thus rejected. Given this, we proceed to Case 5, which constitutes the first

Table 4

Output Equation ML Estimates: Federal Funds Rate Measure, 1948:Q1–1987:Q4

Variable	Case 0	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7
Constant	0.582	0.544	0.581	0.587	0.595	0.571	0.590	0.569
	(0.097)	(0.114)	(0.111)	(0.119)	(0.097)	(0.095)	(0.097)	(0.095)
Δy_{t-1}	0.253	0.301	0.255	0.246	0.246	0.276	0.252	0.279
	(0.081)	(0.098)	(0.096)	(0.098)	(0.081)	(0.081)	(0.081)	(0.080)
Δtbr_t	0.002	0.002	0.002	0.003	0.003	0.003	0.002	0.003
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Δtbr_{t-1}	0.001	0.001	0.001	0.001	0.001	0.000	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.0001)	(0.001)	(0.001)	(0.001)
Δe_t^{S+}	—	–0.304	–0.006	0.000	—	0.000	0.000	—
		(0.295)	(1.413)	(0.237)		(0.082)	(0.102)	
Δe_t^{S-}	—	–0.086	–0.006	0.068	0.067	0.000	0.000	0.460
		(0.145)	(1.413)	(0.113)	(0.081)	(0.082)	(0.102)	(0.213)
Δe_t^{B+}	—	–0.304	–0.006	0.000	—	0.246	—	—
		(0.323)	(1.413)	(0.237)		(0.120)		
Δe_t^{B-}	—	0.694	–0.006	0.068	0.067	0.246	—	—
		(0.289)	(1.413)	(0.113)	(0.081)	(0.120)		
σ_Y	0.850	0.813	0.850	0.847	0.849	0.829	0.851	0.825
$Q(1)$	0.000	0.001	0.001	0.044	0.048	0.064	0.000	0.002
	[0.999]	[0.977]	[0.978]	[0.834]	[0.827]	[0.800]	[0.990]	[0.968]
$Q(10)$	8.747	7.226	8.857	7.662	7.700	7.306	8.762	7.235
	[0.556]	[0.704]	[0.546]	[0.662]	[0.658]	[0.696]	[0.555]	[0.703]
$QQ(1)$	0.384	0.346	0.397	0.194	0.215	0.519	0.406	0.377
	[0.536]	[0.556]	[0.529]	[0.660]	[0.643]	[0.471]	[0.524]	[0.539]
$QQ(10)$	10.25	12.41	10.34	9.950	10.047	12.136	10.325	12.603
	[0.419]	[0.259]	[0.411]	[0.445]	[0.436]	[0.276]	[0.413]	[0.247]
Log likelihood	–300.72	–285.48	–300.71	–291.17	–292.86	–288.99	–293.20	–288.34
LR test		$30.5^{0)}$	$30.5^{1)}$	$11.4^{2)}$	$3.39^{3)}$	$7.01^{4)}$	$8.43^{5)}$	$5.71^{6)}$
		[0.000]	[0.000]	[0.003]	[0.066]	[0.030]	[0.004]	[0.127]

NOTE: See note to Table 1. 0) *LR* test of Case 0 vs. Case 1; 1) *LR* test of Case 2 vs. Case 1; 2) *LR* test of Case 3 vs. Case 1; 3) *LR* test of Case 4 vs. Case 3; 4) *LR* test of Case 5 vs. Case 1; 5) *LR* test of Case 6 vs. Case 5; 6) *LR* test of Case 7 vs. Case 1.

step in testing for the menu-cost type asymmetry. These restrictions are (marginally) rejected since the probability value of the LR test is 3 percent. However, even if one were willing to accept the hypothesis, inspecting Table 4 reveals that the parameter estimates imply that the small shocks are neutral, whereas the big shocks have real effects; menu-cost theories imply the opposite pattern.

However, we still need to look at Case 7, which introduces the restrictions based on the model of Ball and Mankiw (1994). Again, we test this case against Case 1 because it might not be nested in Case 3 and/or Case 5.[26] The likelihood of this specification is higher than the likelihood of Case 5. The LR test implies that we cannot reject the restrictions and the probability value is as high as 13 percent. Furthermore, the coefficient on small negative shocks is now significant at standard significance levels. This result is somewhat surprising, suggesting that the only monetary shocks that have real effects are small negative ones (i.e., contractionary policies)

[26] Given the evidence from Case 1, we also checked if only big negative shocks have real effects. Case 7 turns out to have a much higher likelihood than this alternative case.

and that such contractions of monetary policy lower output. Thus, the empirical evidence seems to be in favor of the hybrid asymmetry. In conclusion, the results indicate very strong empirical evidence in favor of the hybrid asymmetry. In one sense, this result provides evidence in favor of the asymmetry hypothesis and shows that one could increase steady-state output by lowering the variance of the monetary policy shocks. Nevertheless, one has to be careful with the interpretation. If the monetary authority were to make monetary policy more predictable, firms might change their pricing policies; this would affect the range of monetary policies that would have real effects. Thus, it is not clear that the policy implication mentioned above holds in this setting.

SUMMARY AND CONCLUSIONS

Asymmetries in the relationship between real aggregate activity and monetary policy is a phenomenon that can arise under a variety of different assumptions about the economy. The specific version of the asymmetry differs between competing theories, and it is often difficult to test the underlying assumptions directly on macroeconomic data. Furthermore, it is not clear that tests of the assumptions at the household or firm level necessarily carry over to the aggregate level. Since such asymmetric effects in principle can have strong implications not only for the way we think about the macroeconomy, but also for the conduct of economic policy, it thus seems important to empirically examine the evidence on these asymmetries using aggregate data.

In this paper we have focused on the possible asymmetries in the way that different monetary policy shocks affect real aggregate activity. The principal aim of our investigation has been to test indirectly for the asymmetries that may arise in macroeconomic models with menu costs, but the analysis may be thought of more broadly in terms of models with imperfections in goods and labor markets. We highlighted the possible distinctions between different monetary policy shocks that may arise in such models, and we compared these with the traditional Keynesian asymmetry that has been investigated empirically in a number of papers.

In principle, the menu-costs models imply a different type of asymmetric effect than the distinction between positive and negative shocks tested for in previous papers. The most important distinction in basic menu-costs models is between big and small shocks as distinguished either by their size (in a non-stochastic environment) or by their variance (in a stochastic environment). However, with steady-state inflation, there may also be a distinction between positive and negative shocks, but the implied asymmetry is different from the traditional Keynesian asymmetry since the latter does not distinguish shocks by their size.

We developed an empirical framework to distinguish between these competing theories and to test for each of them; we applied this to U.S. postwar data. Our results indicated that, when using M1, the evidence is slightly mixed—since we cannot reject either that shocks are symmetric or that negative shocks have the same effects as positive shocks (but both types of shocks are non-neutral). These results, however, may be hampered by the instability of M1 demand, and we considered the same analysis using the federal funds rate as the monetary policy measure rather than M1. In these data, which we have more faith in, we found very strong evidence in favor of only small negative shocks having real effects. Thus, the U.S. data seem to indicate evidence in favor of the asymmetry implied by menu-costs models in environments with positive steady-state inflation.

It would be interesting to extend this analysis along two lines. First, one might wish to look into other versions of asymmetric effects. One possible direction could be to look into how economic policy affects output in different phases of the business cycle. Another possibility is to look into the effects of nominal demand shocks and their potential asymmetric effects in stochastic dynamic general equilibrium models. We plan to investigate these matters in future research.

REFERENCES

Agénor, Pierre-Richard. "Asymmetric Effects of Monetary Policy Shocks." Working paper, The World Bank, 2001.

Akerlof, George A. and Yellen, Janet L. "A Near-Rational Model of the Business Cycle with Wage and Price Inertia." *Quarterly Journal of Economics*, 1985, *100*(Suppl 4), pp. 823-38.

Baba, Y.; Hendry, David F. and Starr, Ross M. "The Demand for M1 in the USA, 1960-1988." *Review of Economic Studies*, January 1992, *59*(1), pp. 25-61.

Ball, Laurence and Mankiw, N. Gregory. "Asymmetric Price Adjustment and Economic Fluctuations." *Economic Journal*, March 1994, *104*(423), pp. 247-61.

Ball, Laurence and Romer, David. "Are Prices Too Sticky?" *Quarterly Journal of Economics*, August 1989, *104*(3), pp. 507-24.

Ball, Laurence and Romer, David. "Real Rigidities and the Non-Neutrality of Money." *Review of Economic Studies*, April 1990, *57*(2), pp. 183-203.

Barro, Robert J. "Unanticipated Money Growth and Unemployment in the United States." *American Economic Review*, March 1977, *67*(2), pp. 101-15.

Barro, Robert J. "Unanticipated Money, Output, and the Price Level in the United States." *Journal of Political Economy*, August 1978, *86*(4), pp. 549-80.

Barro, Robert J. and Hercowitz, Zvi. "Money Stock Revisions and Unanticipated Money Growth." *Journal of Monetary Economics*, April 1980, *6*(2), pp. 257-67.

Barro, Robert J. and Rush, Mark. "Unanticipated Money and Economic Activity," in Stanley Fischer, ed., *Rational Expectations and Economic Policy*. Chicago: University of Chicago Press, 1980.

Belongia, Michael T. "Measurement Matters: Recent Results from Monetary Economics Reexamined." *Journal of Political Economy*, October 1996, *104*(5), pp. 1065-83.

Bernanke, Ben S. and Gertler, Mark. "Agency Costs, Net Worth, and Business Fluctuations." *American Economic Review*, March 1989, *79*(1), pp. 14-31.

Blanchard, Olivier J. and Kiyotaki, Nobuhiro. "Monopolistic Competition and the Effects of Aggregate Demand." *American Economic Review*, September 1987, *77*(4), pp. 647-66.

Boschen, John F. and Grossman, Herschel I. "Tests of Equilibrium Macroeconomics Using Contemporaneous Monetary Data." *Journal of Monetary Economics*, November 1982, *10*(3), pp. 309-33.

Caballero, Ricardo J. and Engel, Eduardo M.R.A. "Microeconomic Rigidities and Aggregate Price Dynamics." *European Economic Review*, May 1993, *37*(4), pp. 697-711.

Christiano, Lawrence J.; Eichenbaum, Martin and Evans, Charles. "The Effects of Monetary Policy Shocks: Evidence from the Flow of Funds." *Review of Economics and Statistics*, February 1996, *78*(1), pp. 16-34.

Cover, James Peery. "Asymmetric Effects of Positive and Negative Money-Supply Shocks." *Quarterly Journal of Economics*, November 1992, *107*(4), pp. 1261-82.

DeLong, J. Bradford and Summers, Lawrence H. "How Does Macroeconomic Policy Affect Output?" *Brookings Papers on Economic Activity*, 1988, (2), pp. 433-94.

Demery, David. "Asymmetric Effects of Money on Real Output: Evidence for the UK." Unpublished manuscript, University of Bristol, 1993.

Garcia, Rene and Schaller, Huntley. "Are the Effects of Monetary Policy Asymmetric?" Unpublished manuscript, Université de Montréal, 1995.

Gertler, Mark. "Financial Capacity and Output Fluctuations in an Economy with Multi-Period Financial Relationships." *Review of Economic Studies*, July 1992, *59*(3), pp. 455-72.

Greenwald, Bruce C. and Stiglitz, Joseph E. "Financial Market Imperfections and Business Cycles." *Quarterly Journal of Economics*, February 1993, *108*(1), pp. 77-114.

Hamilton, James D. "Rational-Expectations Econometric Analysis of Changes in Regime: An Investigation of the Term Structure of Interest Rates." *Journal of Economic Dynamics and Control*, June/September 1988, *12*(2/3), pp. 385-423.

Hamilton, James D. "A New Approach to the Economic Analysis of Non-Stationary Time Series and the Business Cycle." *Econometrica*, March 1989, *57*(2), pp. 357-84.

Hamilton, James D. "Analysis of Time Series Subject to Changes in Regime." *Journal of Econometrics*, July-August 1990, *45*(1/2), pp. 39-70.

Hamilton, James D. "The Daily Market for Federal Funds." *Journal of Political Economy*, February 1996, *104*(1), pp. 26-56.

Hamilton, James D. *Time Series Analysis*. Princeton, NJ: Princeton University Press, 1994.

Hooker, Mark A. "What Happened to the Oil Price-Macroeconomy Relationship?" *Journal of Monetary Economics*, October 1996, *38*(2), pp. 195-213.

Hooker, Mark A. and Knetter, Michael M. "The Effects of Military Spending on Economic Activity: Evidence from State Procurement Spending." Working Paper, Wellesley College, 1996.

Kandil, Magda. "Asymmetry in the Effects of Monetary and Government Spending Shocks: Contrasting Evidence and Implications." *Economic Inquiry*, April 2002, *40*(2), pp. 288-313.

Karras, Georgios. "Are the Output Effects of Monetary Policy Asymmetric? Evidence from a Sample of European Countries." *Oxford Bulletin of Economics and Statistics*, May 1996, *58*(2), pp. 267-78.

Lo, Ming Chien and Piger, Jeremy. "Is the Response of Output to Monetary Policy Asymmetric? Evidence from a Regime-Switching Coefficients Model." Working Paper 2001-022, Federal Reserve Bank of St. Louis, July 2003.

Lucas, Robert E. Jr. "Expectations and the Neutrality of Money." *Journal of Economic Theory*, April 1972, *4*(2), pp. 103-24.

Lucas, Robert E. Jr. "An Equilibrium Model of the Business Cycle." *Journal of Political Economy*, December 1975, *83*(6), pp. 1113-44.

Macklem, Tiff; Paquet, Alain and Phaneuf, Louis. "Asymmetric Effects of Monetary Policy: Evidence from the Yield Curve." CREFE Working Paper No. 42, Université de Montréal,1996.

Mankiw, N. Gregory. "Small Menu Costs and Large Business Cycles: A Macroeconomic Model of Monopoly." *Quarterly Journal of Economics*, May 1985, *10*(2), pp. 529-38.

Mishkin, Fredric C. "Does Anticipated Policy Matter? An Econometric Investigation." *Journal of Political Economy*, February 1982, *90*(1), pp. 22-51.

Parker, Randall E. and Rothman, Philip. "An Examination of the Asymmetric Effects of Money Supply Shocks in the Pre-World War I and Interwar Periods." Working paper, East Carolina University, 2000.

Phillips, Kerk L. "A Two-Country Model of Stochastic Output with Changes in Regime." *Journal of International Economics*, August 1991, *31*(1-2), pp. 121-42.

Ravn, Morten O. and Sola, Martin. "Stylized Facts and Regime Changes: Are Prices Procyclical?" *Journal of Monetary Economics*, December 1995, *36*(3), pp. 497-526.

Ravn, Morten O. and Sola, Martin "A Reconsideration of the Empirical Evidence on the Asymmetric Effects of Money Supply Shocks: Positive vs. Negative or Big vs. Small?" Unpublished manuscript, University of Aarhus, Denmark, March 1996.

Ravn, Morten O. and Sola, Martin "Business Cycle Dynamics: Predicting Transitions with Macro Variables." Unpublished manuscript, Universitat Pompeu Fabra, 1997.

Senda, Takashi. "Asymmetric Effects of Money Supply Shocks and Trend Inflation." *Journal of Money, Credit, and Banking*, February 2001, *33*(1), pp. 65-89.

Sensier, Marianne. "The Asymmetric Effect of Monetary Policy in the UK." Working paper, University of Oxford, 1996.

Shleifer, Andrei and Vishny, Robert W. "Liquidation Values and Debt Capacity: A Market Equilibrium Approach." *Journal of Finance*, September 1992, *47*(4), pp. 1343-66.

Sola, Martin and Driffill, E. John. "Testing the Term Structure of Interest Rates Using a Stationary Vector Autoregression with Regime Switching." *Journal of Economic Dynamics and Control*, May-July 1994, *18*(3/4), pp. 601-28.

Stock, James H. and Watson, Mark W. "A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems." *Econometrica*, July 1993, *61*(4), pp. 783-820.

Weise, Charles L. "The Asymmetric Effects of Monetary Policy: A Nonlinear Vector Autoregression Approach." *Journal of Money, Credit, and Banking*, February 1999, *33*(1), pp. 85-108.

Appendix

DATA DESCRIPTION

All variables studied in this paper are sampled at the quarterly frequency and were de-seasonalized from the source. The first set of data was kindly supplied by James Peery Cover and is described in detail in Cover (1992). The sample period covers 1948:Q1–1987:Q4. The variables used here are defined as follows:

m1 = the logarithm of M1
y = the logarithm of GDP in constant prices
b = the logarithm of the monetary base
u = the unemployment rate
fs = the logarithm of the federal government's budget surplus
tbr = the T-bill rate

The second set of data corresponds to the data set studied in Christiano, Eichenbaum, and Evans (1996). These data were obtained from the Datastream database. The sample period covers 1959:Q3–1995:Q3. The variables used here are defined as follows:

–ff = the negative of the federal funds rate
y = the logarithm of GDP in constant prices
nbr = the logarithm of the sum of non-borrowed reserves and extended credit
py = the logarithm of the implicit GDP deflator
trs = the logarithm of total reserves
tbr = the T-bill rate

THE FILTER

It is assumed that one of the variables included in the filter is governed by a scalar state variable. The other variable(s) is (are) not allowed to switch and is (are) de-meaned. The filter involves the following five steps.

Step 1. Let y and x be the variables that are observed, and let s be the unobserved state variable. Calculate the density of the m past states and the current state conditional on the information included in y_{t-1}, x_{t-1} and all past values of y and x:

$$\begin{aligned}&p(s_t,s_{t-1},,s_{t-m}\,|\,y_{t-1},y_{t-2},,y_0,x_{t-1},x_{t-2},,x_0)\\&=p(s_t\,|\,s_{t-1})p(s_{t-1},s_{t-2},,s_{t-m}\,|\,y_{t-1},y_{t-2},,y_0,x_{t-1},x_{t-2},,x_0),\end{aligned} \tag{A.1}$$

where $p(s_t|s_{t-1})$ is the transition probability matrix of the states that are assumed to follow a Markov process. As in all subsequent steps, the second term on the right-hand side is known from the preceding step of the filter. In the present case the probability on the left-hand side of equation (A.1) is known from the input to the filter, which in turn represents the result of the iteration at date t–1 (from step 5 described below).

Initial values for the parameters and the initial conditions for the Markov process are required to start the filter. The unconditional distribution, $p(s_m, s_{m-1},, s_0)$, has been chosen for the first observation.

Step 2. Calculate the joint conditional density of y_t and $(s_t, s_{t-1},,s_{t-m})$,

$$\begin{aligned}&p(y_t,s_t,s_{t-1},,s_{t-m}\,|\,y_{t-1},y_{t-2},,y_0,x_{t-1},x_{t-2},,x_0)\\&=p(y_t\,|\,s_t,,s_{t-m},y_{t-1},,y_0,x_{t-1},,x_0)p(s_t,,s_{t-m}\,|\,y_{t-1},,y_0,x_{t-1},,x_0),\end{aligned} \tag{A.2}$$

where we assume that

$$\begin{aligned}&p(y_t,s_t,s_{t-1},,s_{t-m}\,|\,y_{t-1},y_{t-2},,y_0,x_{t-1},x_{t-2},,x_0)\\&=\frac{1}{(2\pi)^{0.5}\left(\sigma_0+\Delta\sigma s_t\right)}\exp\left(-\left(2\left(\sigma_0+\Delta\sigma s_t\right)^2\right)^{-1}u_{st}^2\right),\end{aligned}$$

where

$$u_{st} \equiv y_t - (\mu_0 + \Delta\mu s_t) - \Phi(L)\big(y_t - (\mu_0 + \Delta\mu s_t)\big) - \Theta(L)x_t. \tag{28}$$

It should be noted that $p(y_t|s_t,,s_{t-m},y_{t-1},,y_0,x_{t-1},,x_0)$ involves $(s_t,,s_{t-m})$, which is a vector that can take on 2^{m+1} values.

Step 3. Marginalize the previous joint densities with respect to the states, which give the conditional density from which the (conditional) likelihood function is calculated:

$$p(y_t \big| y_{t-1},,y_0,x_{t-1},,x_0) = \sum_{s_t=0}^{1}\sum_{s_{t-1}=0}^{1}\cdot\cdot\sum_{s_{t-m}=0}^{1} p(y_t,s_t,,s_{t-m}|y_{t-1},,y_0,x_{t-1},,x_0). \tag{A.3}$$

Step 4. Combining the results from steps 2 and 3, calculate the joint density of the state conditional on the observed current and past realizations of y:

$$p(s_t,s_{t-1},,s_{t-m} \big| y_t,,y_0,x_t,,x_0) = \frac{p(y_t,s_t,,s_{t-m}|y_{t-1},,y_0,x_{t-1},,x_0)}{p(y_t|y_{t-1},,y_0,x_{t-1},,x_0)}. \tag{A.4}$$

Step 5. The desired output is then obtained from

$$p(s_t,s_{t-1},,s_{t-m+1} \big| y_t,,y_0,x_t,,x_0) = \sum_{s_{t-m}=0}^{1} p(s_t,s_{t-1},,s_{t-m}|y_t,,y_0,x_t,,x_0). \tag{A.5}$$

The output of step 5 is used as an input to the filter in the next iteration. Estimates of the parameters are calculated by maximizing the sample likelihood, which can be calculated from step 3.

All nonproprietary and nonconfidential data and programs for the articles written by Federal Reserve Bank of St. Louis staff and published in *Review* are available to our readers on our web site: **research.stlouisfed.org/publications/review**. Also, you may request data and programs on either disk or hard copy: **Research Department**, Federal Reserve Bank of St. Louis, P.O. Box 442, St. Louis, MO 63166-0442. Please include the author, title, issue date, and page numbers with your request.

These data and programs are also available through **Inter-university Consortium for Political and Social Research (ICPSR)**. Member institutions may request data through the CDNet Order facility. Nonmembers may write to ICPSR, Institute for Social Research, P.O. Box 1248, Ann Arbor, MI 48106-1248; call 734-998-9900; or e-mail **netmail@icpsr.umich.edu**.

General data can be obtained through **FRED (Federal Reserve Economic Data)**, a database providing U.S. economic and financial data and regional data for the Eighth Federal Reserve District. You may access FRED through our web site: **research.stlouisfed.org/fred**.

ON THE INTERNET
research.stlouisfed.org

THE FEDERAL RESERVE BANK of ST. LOUIS

Federal Reserve Bank of St. Louis
P.O. Box 442
St. Louis, MO 63166-0442